STRUCTURAL AND FUNCTIONAL STUDIES ON BOVINE ß-LACTOGLOBULIN

by

ALAN S. McALPINE

Ph.D.

UNIVERSITY OF EDINBURGH

1991

DECLARATION

I declare that this thesis was composed by me, that the work of which it is a record was done by me, except where stated in the thesis. This work has not been accepted elsewhere in any previous application for a degree. All of the sources of information have been aknowledged.

Since everything that is moved functions as a sort of instrument of the first mover, if there was no first mover, then whatever things are in motion would be simply instruments. Of course, if an infinite series of movers and things moved were possible, with no first mover, then the whole infinity of movers and things moved would be instruments. Now, it is ridiculous, even to unlearned people, to suppose that instruments are moved but not by any principal agent. For, this would be like supposing that the construction of a box or bed could be accomplished by putting a saw or a hatchet to work without any carpenter to use them. Therefore, there must be a first mover existing above all - and this we call God.

From "A Prayer for Owen Meany" by John Irving.

To my Mother and Brother

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Lindsay Sawyer for his help and his guidance throughout the entirety of this project.

I would also like to thank the following people:

-Paul Adams for his help with molecular dynamics and in producing some of the figures,

-Dr. Paul Taylor for his tireless help with the computing throughout ,

-Dr. Sandy Blake for help with small molecule work and guidance with the diffractometer,

-Stella Fawcett for help with the crystallisation work,

-Dr. Gordon Harkiss for allowing me to use his laboratory in Edinburgh Vet school and for his help in the use of monoclonal antibodies,

-Fiona for giving me the support and help when I most needed it (thank you)

-and finally to Mary, Joao, Elspeth, Fred and Pete without whom life in in room 302 would never have been the same.

The financial support of the Science and Engineering Research council is gratefully acknowledged.

Abstract

Structural and functional studies on ß-lactoglobulin

A.S. McAlpine

The existing medium resolution structure of the predominant bovine milk whey protein, ß-lactoglobulin (Blg) showed that the fold of the monomer consisted of a barrel of 9 ß-strands and one 3-turn α-helix. Its similarity to retinol binding protein (RBP), two bilin binding proteins and its close sequence similarity to other transport proteins, has resulted in its inclusion in a superfamily, the lipocalycins. The ß-barrel forms an hydrophobic pocket which is thought responsible for its ability to bind a series of small hydrophobic molecules. The function of the protein is uncertain.

The possibility of engineering the protein's hydrophobic pocket so that it can carry small hydrophobic drug molecules through the stomach, where their presence may be detrimental is investigated in the following manner.

Digestion experiments indicated that the protein has remarkable resistance to bovine pepsin and, to a lesser extent, trypsin. Its resistance to human pepsins is such that it would allow its passage through the stomach without any degradation. The presence of a ligand bound to the protein was shown to enhance the resistance to the protease trypsin.

The solution of two small molecule structures is described and serves as an introduction to the technique of X-ray crystallography. Refinement of the existing model, lattice Y at pH 7.8 (space group $B22_12$, a=55.7 Å, b=67.2 Å, c=81.7 Å) encounters some problems and these are discussed. A new medium resolution data set was collected and allows a more accurate model to be obtained. Refinement with the least squares package TNT gives an R-factor of 20.1% at 3.0 Å. The rescaling of existing high resolution data is described which will be merged with the medium resolution data set.

Crystals of the protein were grown from ammonium sulphate at pH 3.0 (space group $P6_3$; a=b=68.49 Å; c=143.17 Å) and data have been collected. A low pH structure will help investigate the proteins remarkable stability under these conditions. It will also help in the investigation of residues within the hydrophobic pocket which may be genetically engineered. Data have also been collected on crystals grown at the proteins isoelectric point, pH 5.2 (space group $P2_1$; a=72.2 Å; b=67.9 Å; c=36.2 Å; ß=92.0°), and examined by molecular replacement which orientates the lattice Y structure in the cell of the unknown.

Blg is the most antigenic milk protein and the purification of monoclonal antibodies towards Blg was undertaken to allow future crystallisation of a complex between the antigen binding fragment and Blg. Such a complex will reveal an antigenic site and allow investigation of the interaction.

The final chapter discusses the results obtained and concludes that Blg is a good candidate to engineer for the production of a drug carrier.

LIST OF CONTENTS

List of contents contd.

List of contents contd.

LIST OF FIGURES

List of figures contd.

LIST OF TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

Blg	:	ß-lactoglobulin
RBP	:	retinol binding protein
MUP	:	Mouse major urinary protein
B_2MG	:	$ß_2$-microglobulin
HLF	:	Human lactoferrin
α-la	:	α-lactalbumin
Ig	:	Immunoglobulin
FMDV	:	Foot and mouth disease virus
Protein HC	:	Protein heterogenous in charge
Fab	:	Antigen binding fragment
Fc	:	Constant fraction of immunoglobulin
DNA	:	Deoxyribonucleic acid
TXA_2	:	Thromboxane A_2
TXB_2	:	Thromboxane B_2
CD	:	Circular dichroism
ORD	:	Optical rotatory dispersion
NMR	:	Nuclear magnetic resonance
MD	:	Molecular dynamics
ELISA	:	Enzyme-linked immunosorbant assay
HPLC	:	High performance liquid chromotography
FPLC	:	Fast protein liquid chromotography
PAGE	:	Polyacrylamide gel electrophoresis
MIR	:	Multiple isomorphous replacement
FFT	:	Fast fourier transform
K_a	:	Association constant
K_d	:	Dissociation constant

List of Abbreviations contd.

MWt	:	Molecular weight
Å	:	Angstroms
fsec	:	femto second
K	:	Kelvin
R-factor	:	Residual factor
Fo, (obs)	:	Observed structure factor
Fc, (calc)	:	Calculated structure factor
rms	:	Root mean square
M	:	Molar
CDR	:	Complimentarity determining region
SD	:	Standard deviation
SRS	:	Synchrotron radiation source
SDS	:	Sodium dodecyl sulphate
PMSF	:	Phenylmethanesulphonyl fluoride
DEAE	:	Diethylamioethyl
EDTA	:	Ethylenediamine tetra acetic acid
PBS	:	Phosphate buffered saline
HRP	:	Horse raddish peroxidase
AP	:	Alkaline phosphatase
PNPP	:	para-Nitro phenyl phosphate
OPD	:	ortho-Phenylene diamine
APS	:	Ammonium persulphate
AS	:	Ammonium sulphate
DTT	:	Dithiothreitol
PEG	:	Polyethylene glycol
KAPPA	:	Rotation about the non-crystallographic axis
R	:	Σ_{hkl} (|Fo| - |Fc|) / Σ_{hkl} (|Fo|)
R_w	:	$\| \Sigma_w$ (|Fo| - |Fc|)^2 / Σ_w Fo^2 $\|^{1/2}$
S	:	$\| \Sigma_w$ (|Fo| - |Fc|)^2 / (m-n)$^{1/2}$
D	:	Dalton
SDADD & SDFAC	:	Constants used to inflate the standard deviations of all (or only of the stronger) reflections as of equation 4. in section 6.2.6 within the CCP4 program AGROVATA.

CHAPTER 1

Introduction.

Chapter 1.

1.1 Introduction.

Milk is a means by which young mammals obtain nutrition and immunological protection via the mother of the species. Ruminants rely solely on milk to transfer antibodies to their young, a function which is undertaken by the placenta in many other species. The content of milk can vary as a result of the diet and the digestive processes of the lactating animal. Milk also varies according to the time after parturition and from species to species. The content in some species is not so critical since the young soon begin to consume other food.

Bovine milk has become part of an extensive industry and is found as the main constituent in many infant feeds. It is also a major part of mans diet across the world.

Bovine milk can be separated into two parts, the curd and the whey. The curd fraction contains fats and caseins and is commonly made into cheese. The whey is the liquid fraction and is rich in many other proteins. The two principal whey proteins (ß-lactoglobulin and α–lactalbumin) are synthesised by the mammary gland, whereas most others are synthesised elsewhere and are transferred unchanged from the blood.

The most abundant bovine whey protein, ß-lactoglobulin (Blg), is the protein of interest within this thesis. Blg was first isolated by Palmer in 1934 and has since been the subject of many physico-chemical studies. The protein has been extensively reviewed on a number of occasions (Tilley, 1960; McKenzie, 1971; Lyster, 1972), the most recent of which can be found as an appendix (2) to this thesis. Only those

properties of direct relevance to this work are reviewed in this Chapter.

1.2 Isolation of ß-lactoglobulin.

Since its first isolation (Palmer, 1934), Blg has been subject to a wide variety of techniques. Primarily this is due to its ease of purification, which has been frequently described, modified and improved over the years (Aschaffenburg & Drewry, 1957; Armstrong *et al.*, 1967; Monaco *et al.*, 1987). However, despite the method of Monaco *et al.* being apparently the best (crystals grown from this preparation diffracted to a higher resolution than those grown from the Aschaffenburg & Drewry method), the method of Aschaffenburg & Drewry still appears to be that used commercially. Bain & Deutsch (1948) purified both bovine and goat Blg by ethanol fractionation but were unable to obtain crystals of either protein, suggesting that some change had taken place.

1.3 Structural Studies.

In the bovine species Blg exists as a dimer with each monomer being made up of 162 amino acids. Before any three dimensional information on Blg had been fully interpreted, the secondary structure of the protein in solution had been analysed by a variety of techniques. Circular dichroism (CD) indicated that it contained 10% α-helix, 43% anti-parallel ß-sheet and 47% unordered structure (Townend *et al.*, 1967). More recent estimates obtained by Raman spectroscopy (Frushour & Koenig, 1975), optical rotatory dispersion (Timasheff *et al.*, 1983) and prediction algorithms (Creamer *et al.*, 1983) are in good agreement with these early CD values.

Many crystal forms exist and the first to be described at medium resolution was the lattice Y model obtained by Papiz *et al.* (1987) and is of bovine Blg crystallised from

ammonium sulphate at pH 7.8. Values of 7% α-helix and 51% anti-parallel ß-sheet were obtained for the secondary structural elements. The α-helix lies between residues 130-140 and is close to that predicted by Creamer *et al.* (129-143) and Green *et al.* (1978).

The molecule is made up of nine anti-parallel ß-strands and one three turn α-helix. The core of the molecule is made up of eight strands which wrap round to form a hydrophobic pocket. The remaining strand (I, Sawyer *et al.*, 1985) is involved in anti-parallel interactions with its symmetry related strand to form the dimer. (Figure 7.2)

The structure revealed the presence of two disulphide bonds between cysteines 66-160 and 106-119. A free sulphydryl exists at position 121 and the structure shows no evidence of disulphide interchange between residues 119 and 121 with Cys-106, as suggested by McKenzie *et al.* (1972). Any interchange would require significant movement of the ß-sheet strands of the protein molecule.

1.4 Genetics of ß-lactoglobulin.

The gene for bovine Blg encodes for a 178 residue pre-Blg. The 16 amino acid signal peptide is cleaved off, resulting in the mature protein. Totsuka *et al.* (1990) have successfully cloned and expressed the bovine gene in yeast (*Saccharomyces cerevisiae*) in the hope of engineering the protein to characterise the effects of individual residues in various of the proteins properties. The gene for the ovine protein has been characterised (Ali & Clark, 1988) and shows close similarity to the gene structures of rat RBP, mouse major urinary protein (MUP), α_1-acid

glycoprotein and apo-lipoprotein D. The sequence homology of the proteins is low but the arrangement of the gene introns and exons suggests that they are of common origin. The genes of Blg and MUP show the closest homology with introns being in phase and all exons are of a similar size.

In this Department the DNA encoding ovine protein has also been cloned into yeast (*Saccharomyces cerevisiae*) and a few residues have been mutated (Paterson, 1991). Expression is at slightly higher level than in the Japanese work. The model structure for bovine Blg is considered as essentially identical to the ovine structure, since they differ in only 5 amino acids. The system is being used as a method of investigating the role of amino acids within the structure of the protein. The replacement of Cys-119 and Cys-121, independently, by Ser indicates that Cys-119 is necessary to allow complete folding and secretion of the mature protein. Removal of Cys-121 seems to have no obvious effect on the protein in that it appears on SDS-PAGE to be identical to the native. This contradicts the theory of McKenzie *et al.* (1972), who suggest that in the normal population of the protein the disulphide bridge exists in both forms. At the dimer interface interactions occur between Ile-147 and Ile-29 and stacking of the imidazoles of His-146 takes place (Papiz *et al.*, 1986). Replacement of Ile-29 by Asp, therefore, introducing a charge into a hydrophobic region should produce monomers for NMR studies and will help investigations of the monomer-monomer interaction.

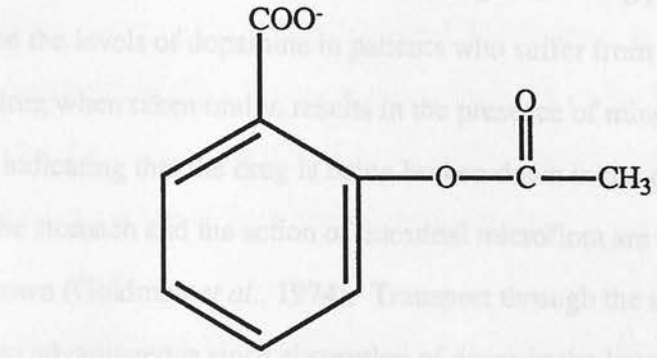
The properties of the hydrophobic pocket also warrant investigation. The binding of small and hydrophobic ligands to Blg suggests that this pocket is responsible. It may even be possible to engineer the pocket allowing it to carry a specific ligand. The most notable of the ligands that are known to bind to Blg is retinol (vitamin A). Its presence within the pocket has been investigated by molecular graphics by Papiz *et al.* (1986). The retinol was placed in an identical position to that which it occupies in

RBP (Newcomer *et al.*, 1984) and shows its tail to be more exposed in the Blg complex. An empty space exists below the retinol, leaving it 10 Å from Trp-19 which lies at the bottom of the calyx. This area is inaccessible to solvent and is blocked in RBP. The retinol can, therefore, fit deeper into the Blg pocket and may explain the higher K_d observed for the Blg-retinol complex ($K_d = 2 \times 10^{-8}$ M; Fugate & Song, 1980) when compared to that with RBP ($K_d = 1.9 \times 10^{-7}$ M; Cogan *et al.*, 1976). Monaco *et al.* (1987) propose a binding site for retinol close to the α-helix, though it is possible that this is an anomalous effect of the crystallisation procedure.

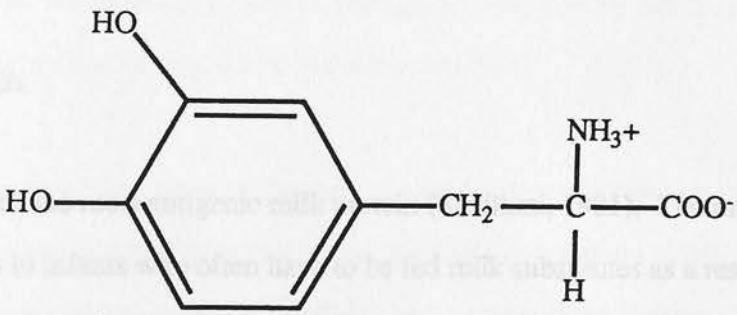
1.5 Proposed use for ß-lactoglobulin.

Helped by the structural knowledge, the successful cloning of the ovine gene and the knowledge that Blg is capable of binding small hydrophobic molecules, it may be possible to engineer the hydrophobic pocket to carry a specific ligand. The type of ligand that might be considered is a drug molecule which at present causes problems when administered to patients who require to take it orally. Binding of a drug molecule to Blg has already been observed. Ellipticine, an anti-tumour drug, binds with a K_d of 1.43×10^{-6} M, though only one molecule binds per dimer (Dodin *et al.*, 1990). The drug is thought to lie at the interface of the two monomers. Problems that can exist in oral administration of a drug are breakdown within the stomach, irritation to the stomach lining and general insolubility of the drug in aqueous media. Therefore, it would be advantagous if a carrier molecule was capable of transporting such molecules through the stomach allowing their absorption in the intestines, especially if this uptake could be promoted by the same molecule.

Possible existing drugs that can be considered in this regard are aspirin and L-dopa (Figure 1.1). They are both of a suitable size and nature that would allow them to sit



(a)



(b)

Figure 1.1 Line drawings of (a) Aspirin (acetyl salacylic acid) and (b) L-Dopa.

in the pocket of Blg. Aspirin is an anti-inflammatory and acts by suppressing the inflammatory effect of prostaglandin. Gastric irritation limits the prescription of high doses of the drug and haemorrhaging can occur. This is a problem to arthritic patients who require to take large amounts of the drug over a long period. L-dopa is used to increase the levels of dopamine in patients who suffer from Parkinsons disease. The drug when taken orally, results in the presence of minor metabolites in patients' urine indicating that the drug is being broken down in the gut. Acid conditions of the stomach and the action of intestinal microflora are the reasons given for this breakdown (Goldman *et al.*, 1974). Transport through the stomach to the intestine is also advantageous since absorption of drugs in the lower gut is more effective than in the stomach (Siurala *et al.*, 1969) .

If Blg is to be considered for use as a drug carrier some of its properties must be examined in relation to its suitability for such a piece of work.

1.6 Antigenicity.

Blg is considered the most antigenic milk protein (Savilhati, 1981). The major problem occurs in infants who often have to be fed milk substitutes as a result of an allergic reaction that occurs towards Blg. Even breast fed infants can show reactions such as colic and diarrhoea as a result of the transfer of bovine milk products from the mothers diet (Jakobsson *et al.*, 1985).

The cross reaction of antibodies raised towards bovine Blg to a factor in human milk originally led to the belief that humans did in fact synthesise a variant of Blg (Liberatori *et al.*, 1979). However, it is now thought that human Blg does not exist and the cross reaction of antibodies is a result of human lactoferrin (Brignon *et al.*,

1985; Monti *et al.*, 1989), β_2-microglobulin (Conti & Godovac-Zimmerman, 1990) or the transfer of bovine Blg from the mothers diet (Axellsson *et al.*, 1986). Sequence comparisons of bovine Blg, human lactoferrin and β_2-microglobulin (Figure 1.2) indicate that a region of Blg between 129-140 (the α-helix) is a possible antigenic site. S-carboxymethylation has revealed four sites on the Blg monomer that may be responsible for antigenicity (Otani *et al.*, 1985). One of these sites agrees with that predicted as a result of the homologous sequences.

In some cases antibodies towards bovine Blg failed to locate the presence of some monomeric Blg's, including that of camel (Kessler & Brew, 1970), unless there was a high antibody titre (Liberatori *et al.*, 1979). This suggests that some differences may exist between the sites recognised in the monomer and the dimer. It is possible that this is a result of the high degree of conservation in the dimers (96-99%) as opposed to that in the monomers (30%). Though, it could also be due to the antigenic determinant stretching across the dimer interface.

It is interesting to note that when antibodies are raised separately towards the A and B variants of Blg, more antibodies are produced towards the A variant (Malik *et al.*, 1988). The higher immune response is likely to be a result of the residues that are unique to the A variant, i.e. residue 64 (Asp and Gly in A and B respectively) or 118 (Val and Ala in A and B respectively). It appears that these stimulate ß-lymphocytes more effectively than those on the B variant, since antigenic determinants appear to be identical. It may also be a result of the greater degree of polymerisation that occurs to the A variant (Townend & Timasheff, 1960).

Monoclonal antibodies towards bovine Blg were made available and in Chapter 4 they are investigated with a view to crystallising the complex between the antigen binding

HLF (54-70) **R A D A V T L D G G F I Y E A G L**

Blg (124-140) **R T P E V D D E A L E K F D K A L**

* * *

HLF (67-76) **E A G L A P Y** - K - **L R**

Blg (130-141) **D E A L E K F D K A L K**

* *

HLF (121-135) **R T A G W N V P I T G L R P F**

Blg (124-136) **R T P** - - **E V D D E A L E K F**

* * * * *

B_2 MG (71-86) - **T P T E K D** - **E Y A C** - **R V N H V T**

Blg (124-140) **R T P** - **E V D D E** - **A L E K F D K A L**

* * * * * *

Figure 1.2. Homologous regions between Blg, B_2 MG and HLF.

HLF = human lactoferrin, B_2 MG = $ß_2$- microglobulin, Blg = ß-lactoglobulin. Identical positions are printed in bold and underscored with *, with conservative changes in bold only. Dashes represent deletions in the sequence.

fragment and the protein. Complexes are generally obtained between the protein and the Fab portion of the antibody since the flexibility of the immunoglobulin molecule around the hinge region, does not allow successful crystallisation. At least three structures of such complexes already exist. Two are of hen egg lysozyme (Amit *et al.*, 1986; Sheriff *et al.*, 1987) and the third is of the influenza virus coat protein, neuraminidase (Colman *et al.*, 1987)

Structure solution of a Fab-Blg complex would allow elucidation of a single antigenic determinant and would also aid understanding of the interaction that occurs between an antigen and an antibody.

1.7 Stability at low pH.

Aschaffenburg and Drewry (1957) took advantage of Blg's stability at low pH during their purification procedure. As yet, however, no-one has given a comprehensive reasoning for the proteins' remarkable acid stability.

Townend and Timasheff (1960) showed that at below pH 4.0 Blg undergoes monomerisation. This dissociation from dimer to monomer is reversible, but is enhanced by lowering the pH (Townend *et al.*, 1960) and by a low protein concentration (McKenzie & Sawyer, 1967). Townend *et al.* (1960) indicate that this dissociation results in the exposure of a hydrophobic surface to the surrounding medium. This agrees with Casal *et al.* (1988) who showed by infrared spectra that when dimerisation occurs at around pH 3.0 a single ß-strand is lost and an anti-parallel ß-sheet is formed. This is a result of dimerisation across the molecular dyad of the I-strands from each monomer to complete the dimer (Papiz *et al.*, 1986).

Investigation of the thermodynamics of folding Blg indicates that the conformational

stability of the folded protein is enhanced at low pH (Kella & Kinsella, 1988). The stability is contributed towards by hydrogen bonds formed from the titration of carboxyl groups. The loss of unfavourable electrostatic interactions of ionised carboxylates upon titration is also a possibility.

1.8 Crystallisation of ß-lactoglobulin.

Investigation of the protein at low pH would be facilitated by a structure of the protein from a crystal which was grown under such conditions. Chapter 5 examines the procedure of growing protein crystals and attempts at growing crystals of Blg under a variety of conditions are examined. Structures from a variety of crystals, grown over a range in pH, will also help evaluate some of the physico-chemical data that are available. Crystals of the lattice X (pH 6.8) and lattice Y (pH 7.8) forms of the protein have already helped in the understanding of the transition (Tanford *et al.*, 1959) which occurs to the protein over the pH range 6 to 8 (Hambling *et al.*, 1987; Yewdall, 1988; Hambling, 1990). The transition leads to the exposure of an anomalous carboxyl in each monomer (Tanford & Taggart, 1961) and comparison of the two structures, suggests that Asp-96 is possibly the anomalous carboxyl (Hambling, 1990).

Consideration for Blg as a drug carrier would also be facilitated by the low pH structure since it would allow examination of the conformation appropriate to what would be encountered within the stomach. Residues can be accurately chosen for mutation and the pocket engineered to allow the binding of a prospective drug molecule. It may in fact be that the pocket closes too tightly at low pH and the removal of residues, to position a ligand, cannot be done with any confidence. It seems unlikely, however, since it has been shown that the protein will bind retinol at

pH 2.0 with a $K_d = 2 \times 10^{-8}$ M (Fugate & Song, 1980). If its natural function is to carry a ligand from mother to neonate then a ligand would have to be bound at low pH to allow the passage of the complex through the stomach. It may be that any drug ligand will have to be introduced to the protein at a higher pH followed by lowering to the required value.

Crystals will also serve as a method of investigating binding ligands to the protein. Various ligands bind to Blg and the crystallisation of a complex with sodium dodecyl sulphate (SDS) is investigated in Chapter 5. SDS has been shown to bind to Blg (McMeekin *et al.*, 1949) and the complex was shown to be more stable than the native form (Groves *et al.*, 1951). Crystals at low pH will serve as a basis for ligand soaking experiments and once the appropriate native structure has been solved then the use of difference electron density maps should pick out the position(s) of the bound molecule.

It is expected that the Blg molecule should be essentially the same in each crystal form and this will allow the technique of molecular replacement to be used. This technique has become more popular in crystallography as a result of protein structures being found which appear to be related on the basis of their sequence similarity. However, it is recommended that the sequence homology between the search molecule and the unknown be quite close and it is preferable if the proteins are known to be related by evolution. Chapter 6 investigates the orientation of a model Blg in the cell of a new crystal form.

1.9 Effects of the Stomach.

Since foodstuffs are ingested in a form that cannot readily be absorbed then they must be

broken down into a smaller form by digestion. Gastric secretion is initiated by a nervous reflex mechanism, resulting in the release of a pale yellow fluid of high acidity, pH around 1.0, containing digestive enzymes which aid in the breakdown of foodstuffs.

If Blg is to be used to carry a ligand/drug to the intestines then it must be stable to the effects of proteolytic digestion. This is examined in Chapter 3 and the effect of a ligand bound are also investigated. Retinol is used as a ligand as a result of its tight binding to the protein.

1.10 Crystallography.

The solution of a high resolution structure is essential in elucidating accurate atomic positions in a protein molecule. At low to medium resolution the overall fold of the protein is evident. A structure of Blg lattice Y, at pH 7.8, has been obtained at 2.8 Å (Sawyer *et al.*, 1985) and high resolution data, to 1.8 Å have been collected also (Hambling, 1990). The aim is to refine the structure to higher resolution by means of the restrained least squares program package TNT (Tronrud *et al.*, 1987) using the data at 1.8 Å. The structure can then be used as an initial means at selecting residues to investigate by genetic engineering. A high resolution structure will also provide a suitable model to search data sets collected from new crystal forms using molecular replacement.

Refinement will be aided by the availability of molecular dynamic programs such as MDXREF (in the GROMOS package; Fujinaga *et al.*, 1989) and XPLOR (Brunger *et al.*, 1987). Molecular dynamics and simulated annealing are becoming more routine in protein structure refinement as a result of their ability to search wider regions of conformational space (Gros *et al.*, 1989; Weis *et al.*, 1989).

Chapter 2 describes the structure solution of two small molecule structures and a overview of the crystallographic process. This serves as an introduction to the technique of X-ray crystallography. Refinement of the lattice Y structure is then covered in Chapter 6 as an extension to the small molecule techniques.

1.11 Thesis Rationale.

The aims of this project were to examine Blg in as comprehensive a way as possible with regards to its suitability as a protein drug carrier. The requirement for such a function is brought about by the general insolubility, lability and toxicity caused by many drugs when administered orally. A method of transporting molecules through the stomach to allow their absorption in the lower gut would therefore be of advantage.

A high resolution structure will allow an accurate description of the protein molecule and explain many of its properties. Further, it will allow investigations as to its validity as a drug carrier.

A comprehensive view of the protein over a range of pH values will be helped by crystallisation of the protein under different conditions which in turn will lead to crystals at a pH relevant to the conditions that would be encountered in the stomach. Chapter 5 describes crystallisation of the protein under different conditions and also of a ligand bound to the protein. Data collected from crystals grown can be examined by techniques such as molecular replacement using the known model as a search molecule (Chapter 6).

The stability of Blg to stomach proteases is of importance if the protein complex is to pass

through the stomach intact and allow release of the ligand in the intestines (Chapter 3). The protein's antigenicity may be of importance to the patients who require to take the drug complex. Monoclonal antibodies were available towards the protein and the activity of these is investigated with a view to the crystallisation of an antigen-anitbody complex (Chapter 4).

CHAPTER 2

Small molecule crystallography.

Chapter 2.

2.1 Introduction.

Our knowledge of the architecture of biological molecules, such as hormones, peptides or macromolecules, is mainly attributed to the technique of X-ray crystallography. Many biological molecules can be crystallised and their precise internal order demonstrated when their crystals are used in a diffraction experiment. Understanding the relationship between the ordered arrangement of the molecules within the crystal and the diffraction pattern obtained allows the elucidation of a molecular structure. Every atom within the unit cell (the crystal's repeating unit) contributes to the intensity of each diffracted observation, as a function of its position and nature. The crystal lattice dimensions influence the angle at which each reflection is observed. A short explanation of how to move from data collected to the solution of the three dimensional crystal structure is required.

2.1.1 Diffraction.

It is easy to understand why X-rays are used when considering the analogy to the light microscope. An object, placed on the stage of a microscope, is illuminated by visible light. The object is normally of such dimensions that it interferes with the visible radiation causing diffraction. The scattered light is caught by the object lens and refocussed through the eyepiece lens and recreates the magnified image for the user. At the atomic level, however, the wavelength of visible light is too large to be disrupted by atoms and a wavelength of radiation comparable to the interatomic distances must be used. Since no known lens system is capable of refocussing X-rays, the process must be done mathematically. The three dimensional crystal lattice amplifies the diffracted reflections allowing them to be detected by either

photographic film or, more commonly nowadays, by a detector system which records reflection intensities electronically.

Conditions for diffraction are described by Braggs Law (Eqn. 1). This can be understood if the crystal is considered as being made up of sets of planes (Figure 2.1).

$$n\lambda = 2\,d\sin\theta \qquad (1)$$

In the equation, λ is the radiation wavelength, n is an integer (the diffraction order), d is the perpendicular spacing between lattice planes within the crystal and θ is the angle the incident beam makes with the crystal plane. This tells us that diffracted rays will only be in phase when the distance travelled by the radiation between adjacent layers is an integral number of wavelengths. Reflections are indexed with the values h, k and l according to the set of planes from which they come. The plane with indices h, k, l intersects the cell axis at a/h, b/k and c/l, where a, b and c are the cell axis lengths.

The relationship between the diffracted ray and the atom positions (with electron density ρ around the atom at position x,y,z) within the unit cell (volume V) is shown below

$$\rho_{(xyz)} = 1/V \sum_{hkl} F_{(hkl)} \exp(-2\pi i(hx+ky+lz)) \qquad (2)$$

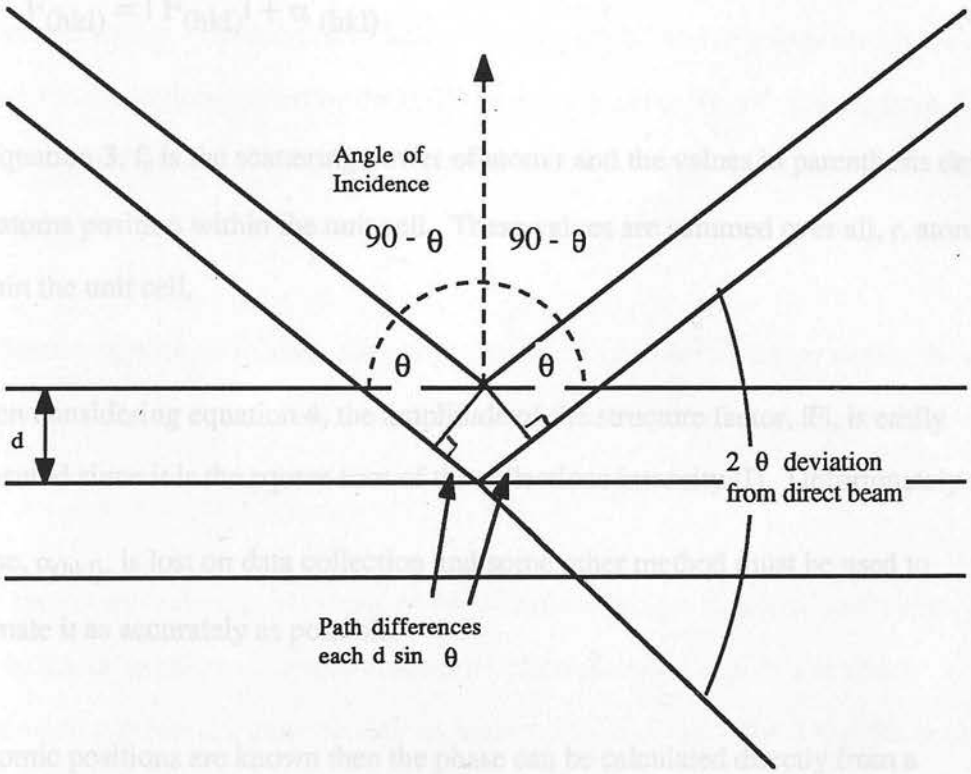


Figure 2.1 Braggs equation is $n\lambda = 2d \sin\theta$. If the path difference of waves scattered by adjacent planes is $2d \sin\theta$ this must equal $n\lambda$ to allow reinforcement of the diffracted beam.

A diffracted ray has a particular combination of amplitude and phase and is commonly known as a structure factor. The structure factor allows the calculation of the electron density within the unit cell, i.e. positions of the atoms within the unit cell.

$$F_{(hkl)} = \Sigma_r f_r \exp 2\pi i (hx_r + ky_r + lz_r) \quad (3)$$

$$F_{(hkl)} = |F_{(hkl)}| + \alpha_{(hkl)} \quad (4)$$

In equation 3, f_r is the scattering power of atom r and the values in parenthesis define the atoms position within the unit cell. These values are summed over all, r, atoms within the unit cell.

When considering equation 4, the amplitude of the structure factor, |F|, is easily measured since it is the square root of the reflections intensity (I). Unfortunately the phase, $\alpha_{(hkl)}$, is lost on data collection and some other method must be used to estimate it as accurately as possible.

If atomic positions are known then the phase can be calculated directly from a rearrangement of equation 3. A trial model allows initial phase estimates which can be refined to give more accurate atomic positions. Some of the methods to obtain a trial structure will be discussed later.

2.1.2 Crystal Characterisation.

Characterisation of a crystal is generally carried out before any data are collected. This involves the determination of the diffraction geometry, i.e. the size and shape of

the unit cell and also the symmetry encompassed therein.

Often information can be derived from crystal morphology though normally investigations begin with some photographic determination of cell constants and space group. The crystal is exposed to the incident beam and the diffracted intensities are recorded on photographic film. This gives a scaled up sampling of diffraction from the unit cell contents. Many types of camera have been constructed to record the reciprocal lattice. The reciprocal lattice has axes a*, b* and c* which are inversely related to the lattice formed by the diffracted X-ray beam, i.e. of dimensions a, b and c. The spatial arrangement of the diffracted beams is a function of the cell dimensions and the wavelength of radiation used.

Symmetry operations relating the position of one molecule to another within the unit cell are usually present. Thus, only the asymmetric unit (the smallest part from which the crystal structure may be obtained) need be considered.

The symmetry is described by one of the 230 space groups (International Tables) and can be identified from absences in the diffraction pattern, i.e. points at which diffraction maxima are expected but no intensity is observed. The 230 different space groups arise from the convolution of the 14 Bravais lattices (possible arrays of points repeating in three dimensions) and the 32 point groups (symmetry operations that leave at least one point, in the object to which they apply, unmoved).

2.1.3 Data Collection.

Data collection involves the recording of diffracted intensities from the crystal and from these intensities the electron density within the crystal can be usually deduced. As has already been discussed, symmetry within the crystal, results in the intensities

of certain reflections being identical. In addition, normally Friedels Law holds, so that the diffraction pattern has a centre of symmetry, i.e. I(hkl) = I(-h-k-l). Thus, only a fraction of the intensities need be collected. However, anomalous scattering results in the breakdown of this law.

At this point it is important to discuss the reciprocal lattice further. The reciprocal lattice has points which are related to planes in real space. Around the crystal an imaginary sphere of radius $1/\lambda$ is drawn (Figure 2.2). This is called the Ewald sphere of reflection. The origin of the reciprocal lattice lies on the edge of this sphere where the incident beam leaves it (point O). Movement of the real lattice also results in a movement of the reciprocal lattice around O. When reciprocal lattice points lie on the surface of the Ewald sphere Braggs law is satisfied and a reflection appears. The amount of unique data that can be obtained is limited by the wavelength (λ) of incident radiation since only the points whose distance from the reciprocal lattice origin is less than the diameter of the sphere will be observed.

The crystal must be moved in such a way that allows the maximum number of reciprocal lattice points to pass through the sphere and the corresponding diffracted rays to be collected. The incident radiation must remain as constant as possible throughout data collection if the intensities of any two reflections are to be compared accurately. Methods of recording intensities involve either photography or some form of electronic counter.

2.1.3.1 Photographic Methods.

There are several photographic techniques for collecting data. The most common method involves the use of the oscillation (or rotation) camera. The crystal is aligned

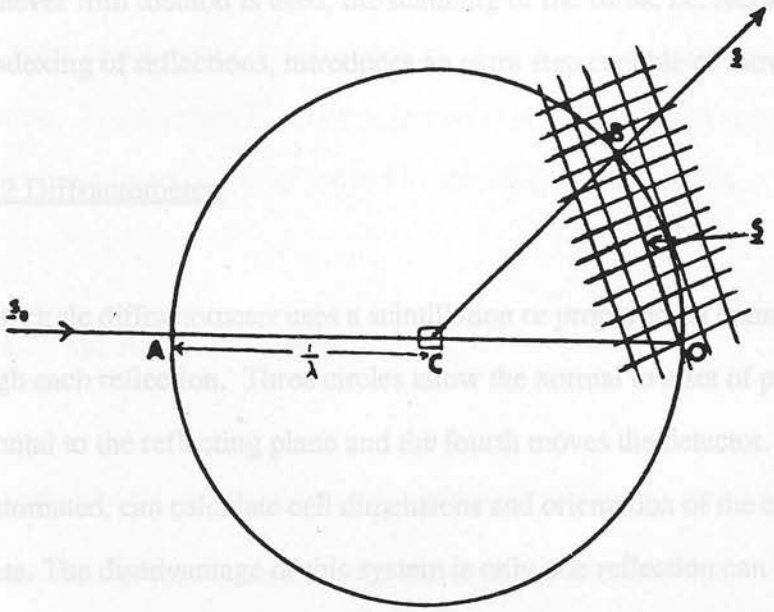


Figure 2.2 The Ewald construction. If point B represents a reciprocal lattice point (hkl), then X-rays will be diffracted in the direction CB. The crystal is rotated in such a waythat the reciprocal lattice point B cuts the sphere of reflection.

with a real axis parallel to the monochromatic (one wavelength) beam. It is then slowly oscillated through a small angle and a small region of the reciprocal lattice recorded on the film. This is continued until all unique regions have been recorded.

A more recent technique, Laue photography, involves polychromatic radiation, which generates a family of Ewald spheres. In a single still all points within these spheres are recorded and a full data set can be recorded in relatively few photographs.

Whichever film method is used, the scanning of the films, i.e. recording intensities and indexing of reflections, introduces an extra step capable of introducing errors.

2.1.3.2 Diffractometer.

A four-circle diffractometer uses a scintillation or proportional counter to step scan through each reflection. Three circles allow the normal to a set of planes to be aligned horizontal to the reflecting plane and the fourth moves the detector. Such instruments are automated, can calculate cell dimensions and orientation of the crystal and collect the data. The disadvantage of this system is only one reflection can be measured at a time. Measuring many reflections is therefore time consuming and can lead to significant radiation damage of the crystal.

2.1.3.3 Area Detectors.

Area detectors help to overcome the problem of radiation damage by measuring many reflections at once. These detectors tend to use the oscillation method but collect a series of images, typically 0.1 - 0.2° rotations.

2.1.3.4 Data Reduction.

Once the raw data for structure solution has been obtained, i.e. intensities, the structure factor modulus (amplitude) must be computed. A relationship exists between |F| and $\sqrt{I}$ which can normally be obtained with reasonable certainty (Eqn. 5)

$$|F_{(hkl)}| = (KI_{(hkl)}/Lp)^{1/2} \quad (5)$$

p, the polarisation factor, is a function of 2θ, and is independent of the method of collection. The constant K is dependent on crystal size and on the beam intensity. L is the Lorentz factor and is described by the following relationship

$$L = \sin\theta/\sin 2\theta\ (\sin^2\theta - \sin^2\mu)^{1/2} \quad (6)$$

The Lorentz factor is dependent on the geometry of the data collection method, with μ being the camera setting angle, and θ the angle the reflection makes with the plane. Any given reflection, or reciprocal lattice point, will take a finite time to pass through the surface of the Ewald sphere, dependent on the scattering angle. The Lorentz factor corrects for this. The effects of absorption and radiation damage are also usually corrected at this stage.

2.1.4 Phase Determination.

Measured data are devoid of phases. If an initial phase estimate can be found then a trial model can be obtained and steadily improved in a cyclic manner (Figure 2.3).

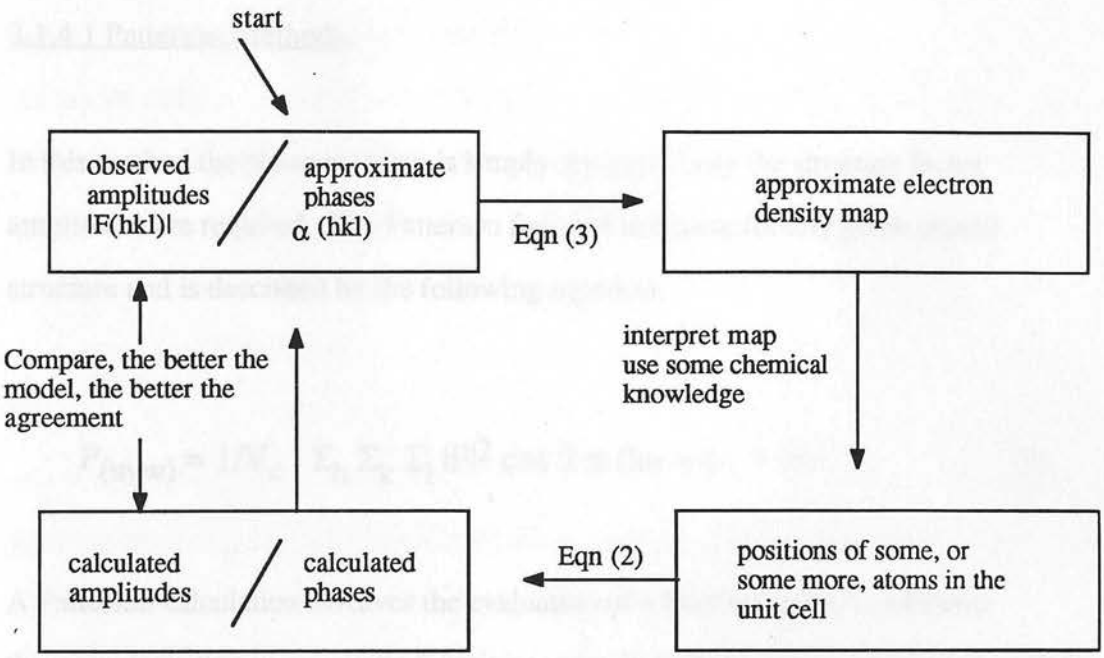


Figure 2.3 Assuming approximate phases can be found, recyclying according to the above scheme can lead to an improvement of the model structure. (Harding, 1991).

Several ways exist for overcoming the phase problem.

2.1.4.1 Patterson Methods.

In this method the phase problem is simply ignored. Only the structure factor amplitudes are required. The Patterson function is unique for any given crystal structure and is described by the following equation:

$$P_{(uvw)} = 1/V_c \quad \Sigma_h \Sigma_k \Sigma_l |F|^2 \cos 2\pi (hu + kv + lw) \qquad (7)$$

A Patterson calculation involves the evaluation of a Fourier series. The Fourier theorem states that any periodic function can be described in terms of sine and cosine. A crystal is periodic in three dimensions, so can be mathematically described by a three dimensional Fourier series.

The Patterson synthesis (Eqn. 7) produces a map whose peaks represent all possible interatomic vectors, all of which start from a common origin. The vector uvw, produces a peak whose height is related to the number of electrons belonging to the atoms at either end of the vector.

Unravelling the map is easier if a small number of atoms within the unit cell have a significantly higher atomic number. Vectors between these atoms dominate the map and a trial structure can be extracted using these atoms alone.

2.1.4.2 Direct Methods.

It can be argued that if a trial structure can be extracted from the peaks within a

Patterson map, then it should be possible to obtain a trial structure from intensities alone. Direct methods operate on this assumption and on the assumption that in both centrosymmetric and non-centrosymmetric structures there is a limit to the possible phases for individual reflections. Electron density cannot be negative throughout the unit cell and will occur in discrete peaks representing atoms. For three reflections that are related in a centrosymmetric structure, their signs are as follows:

$$s(H) = s(K)\, s(H+K) \qquad (8)$$

where s is the sign of, H = hkl, and K = h'k'l'. It is possible to derive the phases for such a relationship and therefore obtain a trial structure if all strong reflections are considered. Similar methods exist for non-centrosymmetric structures.

2.1.4.3 Molecular Replacement.

The term molecular replacement covers a variety of techniques which are designed to make use of the occurrence of a particular molecule in more than one crystallographic environment. For example a particular molecule may be forced to crystallise in a variety of different forms or a given molecule may made up of several identical subunits.

It is commonly used in protein crystallography to generate initial phases for a new model. The calculated diffraction pattern from a known structure is superimposed on that of a new data set to find the best orientation. Once the orientation is located a translational search is undertaken to give the best fit for the model on the unknown data. The phases from the model are then used as starting phases for the new data set. This technique is used in Chapter 6 and is described further there.

2.1.5 Refinement.

Once an approximation of phases has been achieved an electron density map can be calculated. This electron density is interpreted with the help of the knowledge of molecular chemistry, i.e. bond lengths and angles etc. However, the atom positions that are chosen are not very accurate. Density is also difficult to interpret and false density information can occur due to errors in phases. The structure must, therefore, be refined. Atom positions modified in order to obtain a more accurate description of the structure.

Refinement involves adjusting the available parameters (coordinates, temperature factors and site occupancy) until the calculated structure factors best agree with the observed structure amplitudes. Structure factors are calculated from the model by the following equation;

$$F_{calc}(hkl) = \Sigma P_j f_j \exp[-B_j (\sin^2\theta/\lambda^2)] \exp 2\pi i (hx_j + ky_j + lz_j) \quad (9)$$

Where P is the site occupancy of the atom, f is the scattering factor and B its isotropic temperature factor.

The accuracy of a structure is given by its residual or R-factor;

$$R = \Sigma_{hkl} | (|F_o| - |F_c|) / \Sigma_{hkl} (|F_o|) \quad (10)$$

Refinement can take either of two approaches, real or reciprocal space refinement: Real space refinement, or Fourier methods, are commonly used to detect parts of the structure once the heavy atom positions have been found or a partial trial structure is available. It involves the use of electron density maps and adjustment of the atomic parameters for the location of missing or wrongly placed atoms.

The production of an |Fo-Fc| map, the difference map, picks out the differences between the observed data and that created by the model. Missing atoms, areas where there should be electrons, but where there are none input from the model, show up as positive peaks. Atoms wrongly placed in the model are areas with too many electrons and show up as troughs or holes in the map. Atom shifts can be indicated since the bias of the density of a wrongly placed atom will be towards its true position.

The quality of these maps is highly dependent upon phasing. In protein crystallography the method can be automated to allow atom movements throughout the structure, though it is rarely used as such. The maps are generally used nowadays to check on gross structural errors. The more common 2|Fo|-|Fc| map is used in model building between refinement cycles.

Normally small molecule crystallography employs reciprocal space refinement which uses the technique of least squares. Least squares is used to find the best fit to a series of experimental data and has been used for three decades in structure refinement. The best model parameters, i.e. positional coordinates and a temperature factor for each atom, are obtained by minimising the square of the sum of the deviations between the experimental data and the data calculated from the model. It is possible to weight some parameters higher than others if some point is deemed to be nearly correct. It is desirable to have more observations than parameters otherwise

the function is underdetermined and will not converge satisfactorily. Often various constraints are introduced to satisfy known chemical criteria such as bond lengths or angles.

2.2 Small Molecule Structures.

The following section describes the structure elucidation of two small molecules. This was undertaken as an exercise to get familiar with the crystallographic process.

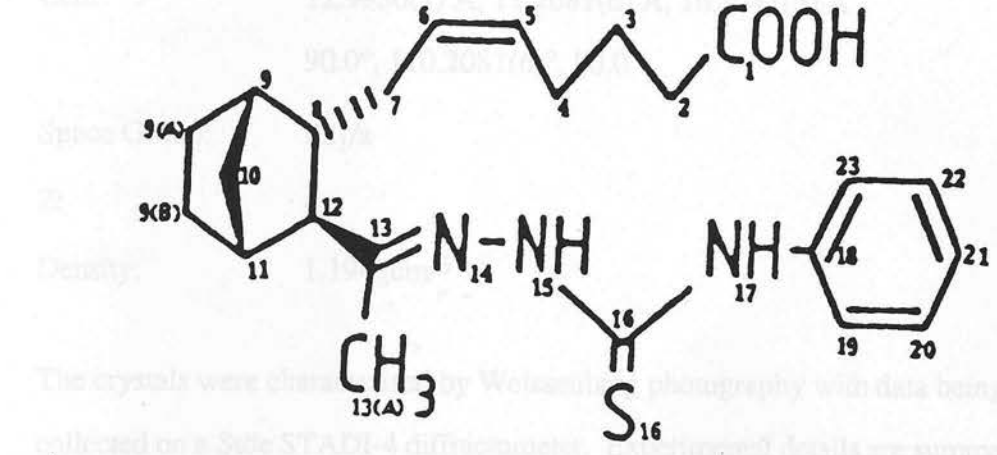
2.3 EP092.

2.3.1. Introduction.

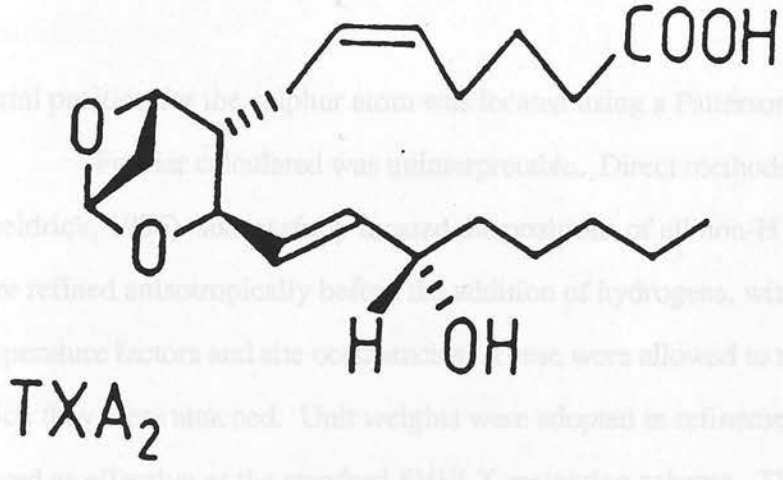
EP092,(+5)-endo-(6'-carboxyhex-2'Z-enyl)-6-exo-{1-[(N-phenylthiocarbamoyl) hydrazono]-ethyl} bicyclo [2.2.1] heptane, is an analogue of the prostaglandins (Wilson & Jones, 1985). It is an antagonist to the thromboxane A_2 (TXA_2) receptor, i.e. it binds tightly to the receptor but fails to elicit an overt response, therefore blocking the action of the the natural agonist. The receptor is involved in vasoconstriction and platelet aggregation. Both EP092 and TXA_2 are shown in figure 2.4.

2.3.2 Experimental.

Crystals (kindly donated by Dr's Wilson, Dawson and Jones from the Department of Pharmacology, University of Edinburgh) were obtained as columns from ethanol at 243 K. The crystal system is described in table 2.1.



EP 092



TXA_2

Figure 2.4 Structures of EP092 and thromboxane A_2

Table 2.1.	Crystal data for EP092
Radiation:	Mo-Kα, $\lambda = 0.71093$ Å, $\mu = 1.55$ cm^{-1}
System:	monoclinic
Formula:	$C_{23}\ H_{31}\ N_3\ O_2\ S$
Cell:	12.9950(7) Å, 11.2081(6) Å, 16.8941(5) Å
	90.0°, 110.2081(6)°, 90.0°
Space Group:	$P2_1/a$
Z:	4
Density:	$1.194 gcm^{-3}$

The crystals were characterised by Weissenberg photography with data being collected on a Stöe STADI-4 diffractometer. Experimental details are summarised in table 2.2. Data was collected with 3 reflections being measured every 2 hours to check for crystal decay or movement.

A trial position for the sulphur atom was located using a Patterson map, though the Fourier calculated was uninterpretable. Direct methods using SHELXS86 (Sheldrick, 1986) successfully located the positions of all non-H atoms. Atoms were refined anisotropically before the addition of hydrogens, with fixed isotropic temperature factors and site occupancies. These were allowed to ride on the atoms to which they were attached. Unit weights were adopted in refinement since they proved as effective as the standard SHELX weighting scheme. The final R-factor achieved was 0.0472.

The program CALC (Gould and Taylor, 1983) was used to provide molecular geometry data presented in table 2.3. Diagrams were prepared using PLUTO (Motherwell, 1972). Final atomic parameters are given in table 2.4.

Table 2.2. Experimental data for EP092.

Crystal dimensions (mm^3):		0.5 x 0.45 x 0.9
Reflections to determine lattice parameters:		41
(2 θ scan range)		(32° - 33°)
Max. sin θ/λ ($Å^{-1}$):		0.538
hkl range;	h:	-13 to 13
	k:	0 to 12
	l:	0 to 17
standard reflections:		-8, 1, 11; -5, 6, 9; 2, 8, 4
total data measured:		3198
data used in refinement:		2778
parameters refined:		265
R:		0.0472
R_w:		0.0434
S:		0.826
max. Δ/σ, last cycle:		0.036
final diff. map ($eÅ^{-3}$):		+ 0.2088
		- 0.2788

Table 2.3 (a) Bond Lengths (Å) with standard deviations

C(1) - O(1)	1.197(4)	C(11) -C(12)	1.558(4)
C(1) -O(1A)	1.310(4)	C(12) -C(13)	1.514(4)
C(1) - C(2)	1.497(5)	C(13) -C(13A)	1.495(4)
C(2) - C(3)	1.500(5)	C(13) -N(14)	1.276(3)
C(3) - C(4)	1.527(5)	N(14) -N(15)	1.395(3)
C(4) - C(5)	1.488(5)	N(15) -C(16)	1.351(4)
C(5) - C(6)	1.313(5)	C(16) -S(16)	1.681(3)
C(6) - C(7)	1.499(4)	C(16) -N(17)	1.342(4)
C(7) - C(8)	1.537(4)	N(17) -C(18)	1.414(4)
C(8) - C(9)	1.538(4)	C(18) -C(19)	1.380(4)
C(8) -C(12)	1.550(4)	C(18) -C(23)	1.387(4)
C(9) -C(9A)	1.530(4)	C(19) -C(20)	1.381(5)
C(9) -C(10)	1.528(4)	C(20) -C(21)	1.362(5)
C(9A) -C(9B)	1.558(5)	C(21) -C(22)	1.372(5)
C(9B) -C(11)	1.538(4)	C(22) -C(23)	1.390(5)
C(10) -C(11)	1.528(4)		

Table 2.3 (b) Angles (degrees) with standard deviations

O(1) - C(1) -O(1A)	122.6(3)	C(8) -C(12) -C(11)	102.55(20)
O(1) - C(1) - C(2)	124.7(3)	C(8) -C(12) -C(13)	116.88(21)
O(1A) - C(1) - C(2)	112.7(3)	C(11) -C(12) -C(13)	111.28(21)
C(1) - C(2) - C(3)	114.4(3)	C(12) -C(13) -C(13A)	116.31(23)
C(2) - C(3) - C(4)	112.4(3)	C(12) -C(13) -N(14)	118.05(23)
C(3) - C(4) - C(5)	112.1(3)	C(13A)-C(13) -N(14)	125.62(25)
C(4) - C(5) - C(6)	128.2(3)	C(13) -N(14) -N(15)	118.22(21)
C(5) - C(6) - C(7)	127.6(3)	N(14) -N(15) -C(16)	117.45(21)
C(6) - C(7) - C(8)	111.18(23)	N(15) -C(16) -S(16)	120.36(20)
C(7) - C(8) - C(9)	116.10(23)	N(15) -C(16) -N(17)	114.70(24)
C(7) - C(8) -C(12)	113.20(21)	S(16) -C(16) -N(17)	124.91(21)
C(9) - C(8) -C(12)	103.10(21)	C(16) -N(17) -C(18)	130.20(23)
C(8) - C(9) -C(9A)	111.48(24)	N(17) -C(18) -C(19)	117.14(25)
C(8) - C(9) -C(10)	100.21(23)	N(17) -C(18) -C(23)	122.85(25)
C(9A) - C(9) -C(10)	101.67(24)	C(19) -C(18) -C(23)	119.9(3)
C(9) -C(9A) -C(9B)	103.08(24)	C(18) -C(19) -C(20)	120.2(3)
C(9A) -C(9B) -C(11)	102.69(24)	C(19) -C(20) -C(21)	120.3(3)
C(9) -C(10) -C(11)	94.44(23)	C(20) -C(21) -C(22)	119.9(3)
C(9B) -C(11) -C(10	101.84(23)	C(21) -C(22) -C(23)	120.9(3)
C(9B) -C(11) -C(12)	107.14(23)	C(18) -C(23) -C(22)	118.7(3)
C(10) -C(11) -C(12)	101.92(22)		

Table 2.3 (c) Geometry around hydrogen bonds (Distances in Å, Angles in degrees).

O(1A) - S(16)	3.147(3)	C(1) -O(1) -H(151)	132.76(24)
H(1A1)- S(16)	2.26(3)	O(1) -H(151) -N(15)	169.90(22)
O(1) - N(15)	2.989(3)	O(1A) -H(1A1) -S(16)	168.6 (23)
O(1) - H(151)	1.919(3)	H(1A1) -S(16) -C(16)	109.0 (7)

Table 2.3 (d). Selected Torsion angles (degrees) with standard deviations

C(9B) -C(11) -C(12) -C(13)	158.67(23)	C(10) -C(11) -C(12) -C(13)	-94.80(25)
C(8) -C(12) -C(13) -C(13A)	158.93(24)	C(8) -C(12) -C(13) -N(14)	-22.4(3)
C(11) -C(12) -C(13) -C(13A)	-83.8(3)	C(6) - C(7) - C(8) - C(9)	-166.96(24)
C(11) -C(12) -C(13) -N(14)	94.9(3)	C(6) - C(7) - C(8) -C(12)	74.1(3)
C(12) -C(13) -N(14) -N(15)	-178.53(21)	C(7) - C(8) - C(9) -C(9A)	-57.8(3)
C(13A)-C(13) -N(14) -N(15)	0.0(4)	C(7) - C(8) - C(9) -C(10)	-164.72(23)
C(13) -N(14) -N(15) -C(16)	178.65(24)	N(14) -N(15) -C(16) -S(16)	177.01(18)
N(14) -N(15) -C(16) -N(17)	-1.2(3)	C(7) - C(8) -C(12) -C(11)	131.94(23)
N(15) -C(16) -N(17) -C(18)	-178.21(25)	C(7) - C(8) -C(12) -C(13)	-106.1(3)
S(16) -C(16) -N(17) -C(18)	3.7(4)	C(16) -N(17) -C(18) -C(23)	36.5(4)

Table 2.4 Fractional Coordinates of Atoms with Standard Deviations

$$Ueq = 1/3 \sum_i \sum_j U_{ij} a^*_i a^*_j a_i . a_j .$$

	x	y	z	Ueq
C(1)	0.93623(25)	0.8608(3)	-0.11108(18)	0.0686(21)
O(1)	0.96601(20)	0.76023(19)	-0.11369(15)	0.0976(19)
O(1A)	0.96605(24)	0.94913(21)	-0.14913(18)	0.1181(23)
C(2)	0.86217(25)	0.9001(3)	-0.06505(19)	0.0730(21)
C(3)	0.8319(3)	0.8032(3)	-0.01590(19)	0.0754(22)
C(4)	0.7564(3)	0.8477(3)	0.02980(19)	0.0788(23)
C(5)	0.71328(25)	0.7483(3)	0.06771(21)	0.0778(24)
C(6)	0.73448(23)	0.7250(3)	0.14805(20)	0.0676(21)
C(7)	0.80924(22)	0.79234(23)	0.22276(17)	0.0616(19)
C(8)	0.89218(21)	0.70838(22)	0.28517(16)	0.0552(17)
C(9)	0.95602(25)	0.76060(24)	0.37282(17)	0.0679(20)
C(9A)	1.0200(3)	0.87260(25)	0.36632(19)	0.0724(21)
C(9B)	1.1132(3)	0.82250(25)	0.33732(20)	0.0738(22)
C(10)	1.04827(24)	0.6687(3)	0.40503(18)	0.0690(20)
C(11)	1.09069(22)	0.68749(24)	0.33220(17)	0.0611(19)
C(12)	0.98492(20)	0.66579(21)	0.25369(16)	0.0522(17)
C(13)	0.97873(20)	0.53800(22)	0.22329(16)	0.0516(17)
C(13A)	1.04104(25)	0.5098(3)	0.16604(20)	0.0791(23)
N(14)	0.92345(17)	0.46360(17)	0.24950(13)	0.0549(14)
N(15)	0.91968(17)	0.34518(17)	0.22320(14)	0.0563(15)
C(16)	0.86267(21)	0.26705(22)	0.25296(17)	0.0539(17)
S(16)	0.84902(7)	0.12425(6)	0.22061(5)	0.0684(5)
N(17)	0.81777(18)	0.31414(19)	0.30642(14)	0.0652(16)
C(18)	0.75602(22)	0.25938(23)	0.35081(16)	0.0564(18)
C(19)	0.67541(24)	0.3273(3)	0.36499(18)	0.0716(21)
C(20)	0.6159(3)	0.2816(4)	0.41169(21)	0.086(3)
C(21)	0.6349(3)	0.1684(4)	0.44292(20)	0.086(3)
C(22)	0.7159(3)	0.1008(3)	0.43010(19)	0.0805(24)
C(23)	0.77782(25)	0.14514(25)	0.38406(18)	0.0684(20)

2.3.3 Discussion.

The crystal structure of EP092 is shown in Figure 2.5. Table 2.3 lists bond lengths, bond angles and torsion angles for the structure. The atoms in the thiocarbamoyl arm are found to be planar with the nitrogen having the greatest deviation from the plane (0.025 Å). The rms deviation from the plane was 0.015 Å. A stereo diagram of the molecular packing of EP092 (Figure 2.6) shows the benzene ring approaching the bicycloheptane moiety. The molecules arrange as dimers across an inversion centre and are stabilised by hydrophobic interactions and by 2 hydrogen bonds. The hydrogen bonds occur by donation of a hydrogen from the N (15) to the carboxyl oxygen, O (1), and of the H atom on O (1A) to the S (16). Distances and angles (Table 2.3) are in agreement with tabulated values (Allen *et al.*, 1987)

EP092 was compared with the thromboxane B_2 (TXB_2, Fortier *et al.*, 1980). TXB_2 differs from TXA_2 by the addition of a water molecule to the anomeric carbon, therefore, converting the acetal of the TXA_2 to a hemiacetal. The torsion around the C4-C5 bond differs by 164°. Superposition of the C7-C8-C12 indicates a torsion change of 90° around the C7-C8 bond. The torsion around C12-C13 differs by 110° allowing the EP092 methyl group to superimpose on the TXB_2 arm. The differences noted seem to be a result of the different ring structures although crystal packing may also have an effect. A comparison of 2 bridged rings in 1-bromo-3-[3-(2-methylene cyclo hex-5-en-1-yl)bicyclo[2.2.1]-hept-2-yl benzene (Bocelli *et al.*, 1984) with that in EP092 (atoms C8 to C12) showed the structures to be fairly rigid with no significant differences. Recently the structure of 4(Z)-6-{[2RS,4RS,5SR] -2- (2-Chlorophenyl)-4-(2-hydroxyphenyl)-1,3-dioxan-5-yl}hex-4-enoic acid (ICI 192605), another thromboxane antagonist, has been reported (Brown & Foubister, 1990). Brown and Foubister indicate the presence of a hydrogen bond in 1,3

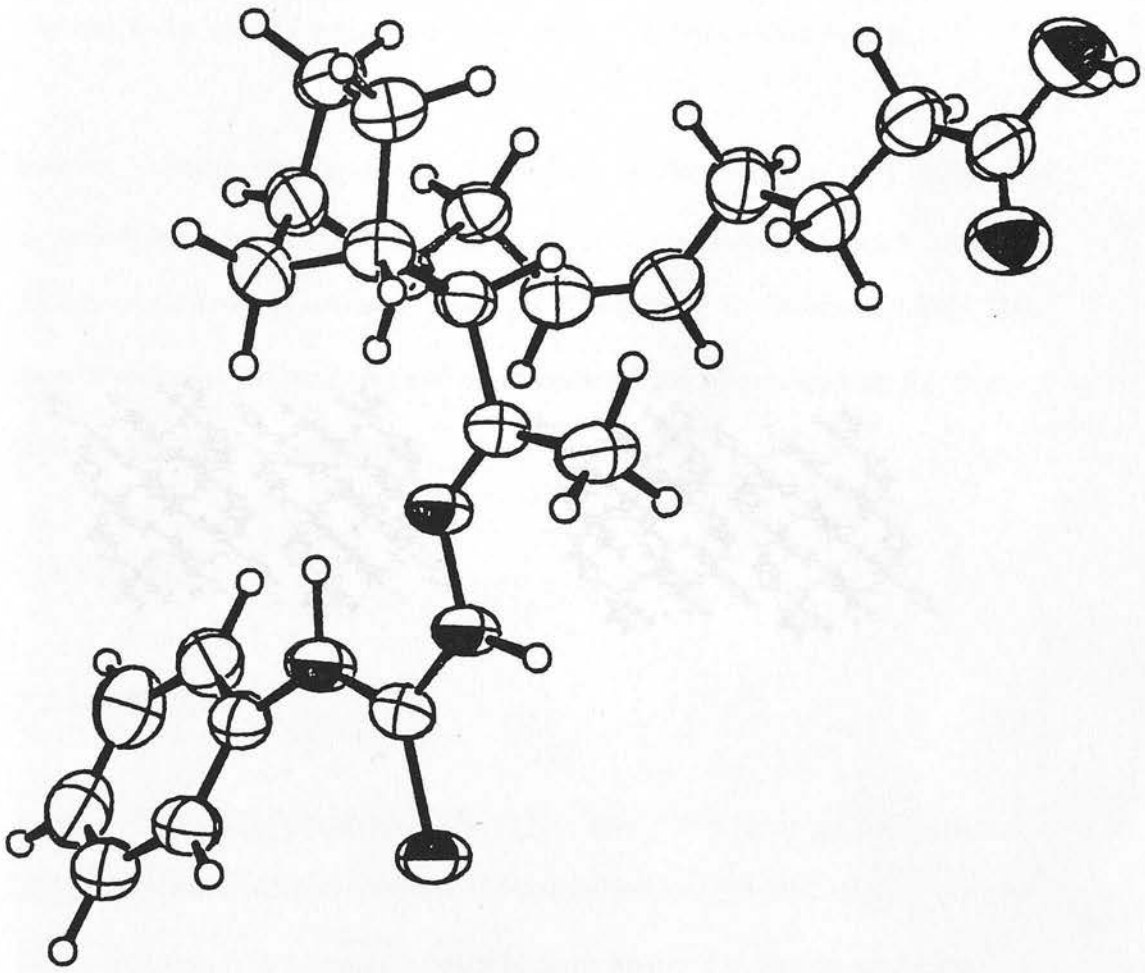


Figure 2.5 Crystal structure of EP092. Non-hydrogen atoms shown as 50 % probability thermal ellipsoids.

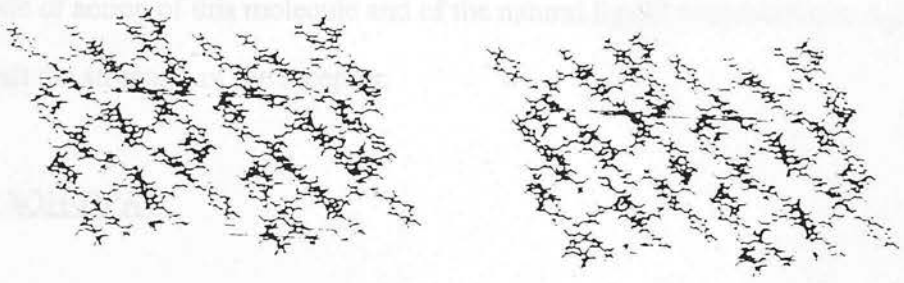


Figure 2.6 Stereo packing of EP092, b-axis projection.

dioxanes may be of importance in receptor recognition. The arms of the hexenoic (from ICI 192605) and heptanoic (from EP092) acid side chains lie in similar though not identical orientations. As a result of these two arms showing similar conformations it may be likely that these arms lie in the receptor pocket . The hydrogen bond in ICI 192605 and the bridged ring in EP092 may both play similar roles, holding the other groups in such a conformation as to allow tighter binding. This may be by specific interaction between the groups and the receptor.

From this structure, and its comparison to the thromboxane B_2 and ICI 192605 it is not entirely clear why EP092 has such an effect on the receptor to which it binds. The torsion differences indicated confer no extra stability to the free molecule. The mode of action of this molecule and of the natural ligand (thromboxane A_2) must await the structure of the receptor.

2.4 8OH-DPAT

2.4.1 Introduction.

8-hydroxy-2-(n-dipropylamino)tetralin (Kirby *et al.*, 1991) is an agonist (a ligand binding tightly to a receptor to *promote* a desired response) acting selectively on the 5-HT_{1A} receptor. The receptor is believed to be involved in thermoregulation, adenylate cyclase modulation and in the control of blood pressure (Peroutka, 1988).

2.4.2 Experimental.

One crystal (kindly donated by Dr's Wilson, Dawson and Jones in the Department of Pharmacology, University of Edinburgh) was grown by isothermal distillation of

ether into methanol. Further attempts to grow crystals proved unsuccessful. The laminar crystal was characterised by Weissenberg photography, and data collected on a Stoe STADI-4 diffractometer using Cu-Kα radiation. Table 2.5 describes the crystal data for 8OH-DPAT.

Table 2.5. Crystal data for 8OH-DPAT.

Radiation:	Cu-Kα, $\lambda = 1.54184$ Å, $\mu = 19.00 cm^{-1}$
System:	monoclinic
Formula:	$C_{16} H_{26} NOCl$
Cell:	9.2229(4) Å; 8.8051(6) Å; 17.6475(7) Å
	90.0°; 94.537(6)°; 90.0°
Space Group:	$P2_1/n$
Z:	4
Density:	$1.151 gcm^{-3}$

An approximate orientation matrix was determined from 8 reflections between 31° and 34° in 2 θ. The three intensity controls showed no variation over the data collection period, although bad peak shapes indicated crystal damage. Further experimental details are layed out in table 2.6.

A trial position for the chlorine atom was determined by Patterson synthesis, though the Fourier was uninterpretable. All non hydrogen atoms were located by direct methods, using SHELXS86 (Sheldrick, 1986) with the exception of the dipropyl amino moiety. All atoms were refined anisotropically by least squares and half occupancies were attributed to atoms C(14)-C(16) but failed to help reduce the

Table 2.6. Experimental data for 8OH-DPAT.

Crystal dimensions (mm^3):		0.43 x 0.18 x 0.33
Reflections to determine lattice parameters:		33
(2 θ scan range)		(31° - 34°)
Max. sin θ/λ ($Å^{-1}$):		0.5567
hkl range;	h:	-11 to 11
	k:	0 to 11
	l:	0 to 13
standard reflections:		-2, 0, 4; -3, 2, 1; 1, 2, 4
total data measured:		2142
data used in refinement:		1550
parameters refined:		144
R:		0.1781
R_w:		-
S:		2.852
max. Δ/σ, last cycle:		1.654
final diff. map ($eÅ^{-3}$):		+ 0.882
		- 0.637

R-factor. The data collection output indicates a possible twinned or split crystal accounting for the bad refinement and high delta/sigma max. A final R-factor of 0.1781 was achieved.

The molecular geometry program CALC was used to calculate bond lengths, bond angles and torsions. Selected angles and torsions are presented in table 2.7. Diagrams were prepared using PLUTO (Motherwell, 1972) and an interactive version of ORTEP (Johnson, 1965; Mallinson & Muir, 1985). Fractional coordinates can be seen in table 2.8.

2.4.3 Discussion.

Bond lengths for the structure are as expected. The crystal structure can be seen in figure 2.7. The molecular packing of the molecule, shown in figure 2.8, indicates few close contacts between the molecules. They lie in an infinite stack, with pairs of molecules being related by an inversion centre. The propyl arms are badly disordered and hinder the search for close contacts. It is possible that a salt bridge between the chlorine and the nitrogen may occur, as is seen in a similar crystal structure, 8MeO-PAT (Kirby *et al.*, 1991), though the distances here are probably a little too long. The C(11)-C(13) propyl arm sits closer to the chlorine than the C(14)-C(16) chain, with the hydrogens on C(11) coming within 3 Å of the chlorine.

A comparison of this crystal structure was made with two related molecules, 8MeO-PAT (Kirby *et al.*, 1991) and 5OH-DPAT (Giesecke, 1980). Superposition of the aromatic rings of 8OH-DPAT and 5OH-DPAT showed differing nitrogen positions resulting in a displacement of the propyl side chains. The angle through the nitrogen from C(11) to C(14) differs by 12°, though this may be a result of the poor model. Superimposition with 8MeO-PAT allows identical positions for the aromatic

Table 2.7 (a) Selected angles (degrees) with standard deviations for 8-OH DPAT. Alternative atom positions are denoted by a prime.

C(2) - N(1) -C(11)	121.4(19)
C(2) - N(1) -C(11')	93.7(26)
C(2) - N(1) -C(14)	107.0(21)
C(11) - N(1) -C(14)	124.9(24)
C(11')- N(1) -C(14)	114.7(29)
C(2) - C(1) - C(9)	114.5(19)
N(1) - C(2) - C(1)	110.3(18)
N(1) - C(2) - C(3)	113.3(18)
C(1) - C(2) - C(3)	110.7(19)
C(2) - C(3) - C(4)	112.0(20)
C(3) - C(4) -C(10)	112.8(19)
O(1) - C(8) - C(7)	122.5(17)
O(1) - C(8) - C(9)	117.5(16)
C(1) - C(9) - C(8)	116.3(17)
C(1) - C(9) -C(10)	123.6(17)
C(4) -C(10) - C(5)	121.1(17)
C(4) -C(10) - C(9)	118.9(17)
N(1) -C(11) -C(12)	110.6(25)
C(11) -C(12) -C(13)	122.0(29)
N(1) -C(14) -C(15)	106.6(29)
C(14) -C(15) -C(16)	102.1(31)
C(14')-C(15')-C(16')	163.3(55)

Table 2.7 (b) Selected Torsion angles (degrees) with standard deviations for 8-OH DPAT. Alternative atom positions are denoted by a prime.

C(11) - N(1) - C(2) - C(1)	-159.2(22)
C(11) - N(1) - C(2) - C(3)	-34.4(29)
C(11')- N(1) - C(2) - C(1)	165.4(27)
C(11')- N(1) - C(2) - C(3)	-69.8(29)
C(14) - N(1) - C(2) - C(1)	48.2(26)
C(14) - N(1) - C(2) - C(3)	173.0(23)
C(2) - N(1) -C(11) -C(12)	-56.1(32)
C(14) - N(1) -C(11) -C(12)	91.4(34)
C(2) - N(1) -C(14) -C(15)	173.0(24)
C(11) - N(1) -C(14) -C(15)	21.6(41)
C(11')- N(1) -C(14) -C(15)	70.6(39)
C(9) - C(1) - C(2) - N(1)	161.9(18)
C(9) - C(1) - C(2) - C(3)	35.6(26)
C(2) - C(1) - C(9) - C(8)	175.5(17)
C(2) - C(1) - C(9) -C(10)	-9.1(29)
N(1) - C(2) - C(3) - C(4)	177.0(18)
C(1) - C(2) - C(3) - C(4)	-58.4(25)
C(2) - C(3) - C(4) -C(10)	54.5(26)
C(3) - C(4) -C(10) - C(5)	153.8(19)
C(3) - C(4) -C(10) - C(9)	-26.4(27)
C(6) - C(5) -C(10) - C(4)	179.8(18)
C(6) - C(7) - C(8) - O(1)	178.0(17)
O(1) - C(8) - C(9) - C(1)	-2.4(26)
O(1) - C(8) - C(9) -C(10)	-178.1(17)
C(7) - C(8) - C(9) - C(1)	175.6(18)
C(1) - C(9) -C(10) - C(4)	4.9(29)
C(1) - C(9) -C(10) - C(5)	-175.3(18)
C(8) - C(9) -C(10) - C(4)	-179.8(18)
N(1) -C(11) -C(12) -C(13)	-178.7(27)

Table 2.8 Fractional coordinates of atoms with standard deviations for 8-OH DPAT. Alternative atom positions are denoted by a prime.

$$U_{eq} = 1/3 \, \Sigma_i \, \Sigma_j \, U_{ij} \, a^*_i \, a^*_j \, a_i . \, a_j .$$

	x	y	z	Ueq
Cl(1)	0.0662(3)	0.19459(24)	0.04771(24)	0.0557(19)
N(1)	0.2931(15)	0.3392(15)	0.1591(14)	0.131(14)
O(1)	0.4566(16)	0.3621(13)	-0.2217(15)	0.172(15)
C(1)	0.3558(21)	0.3960(18)	-0.0263(22)	0.155(21)
C(2)	0.3058(21)	0.4272(17)	0.0876(18)	0.129(18)
C(3)	0.1822(22)	0.4847(17)	0.0720(21)	0.141(20)
C(4)	0.2004(20)	0.5766(20)	0.0038(19)	0.144(20)
C(5)	0.2284(14)	0.6174(11)	-0.1947(19)	0.142(20)
C(6)	0.2803(14)	0.5956(11)	-0.2951(19)	0.138(20)
C(7)	0.3580(14)	0.5111(11)	-0.3054(19)	0.143(21)
C(8)	0.3838(14)	0.4484(11)	-0.2153(19)	0.138(21)
C(9)	0.3319(14)	0.4703(11)	-0.1148(19)	0.119(17)
C(10)	0.2542(14)	0.5548(11)	-0.1045(19)	0.136(21)
C(11)	0.207(3)	0.341(3)	0.2531(24)	0.145(5)
C(11')	0.284(5)	0.397(4)	0.259(5)	0.145(5)
C(12)	0.249(4)	0.426(3)	0.334(3)	0.145(5)
C(12')	0.150(5)	0.350(4)	0.323(5)	0.145(5)
C(13)	0.180(4)	0.444(3)	0.440(3)	0.145(5)
C(13')	0.140(5)	0.380(4)	0.418(5)	0.145(5)
C(14)	0.424(4)	0.280(3)	0.156(4)	0.145(5)
C(14')	0.913(6)	0.219(5)	-0.302(5)	0.145(5)
C(15)	0.417(4)	0.191(3)	0.246(4)	0.145(5)
C(15')	0.929(5)	0.324(5)	-0.316(5)	0.145(5)
C(16)	0.548(3)	0.138(3)	0.234(3)	0.145(5)
C(16')	0.918(5)	0.434(4)	-0.305(4)	0.145(5)

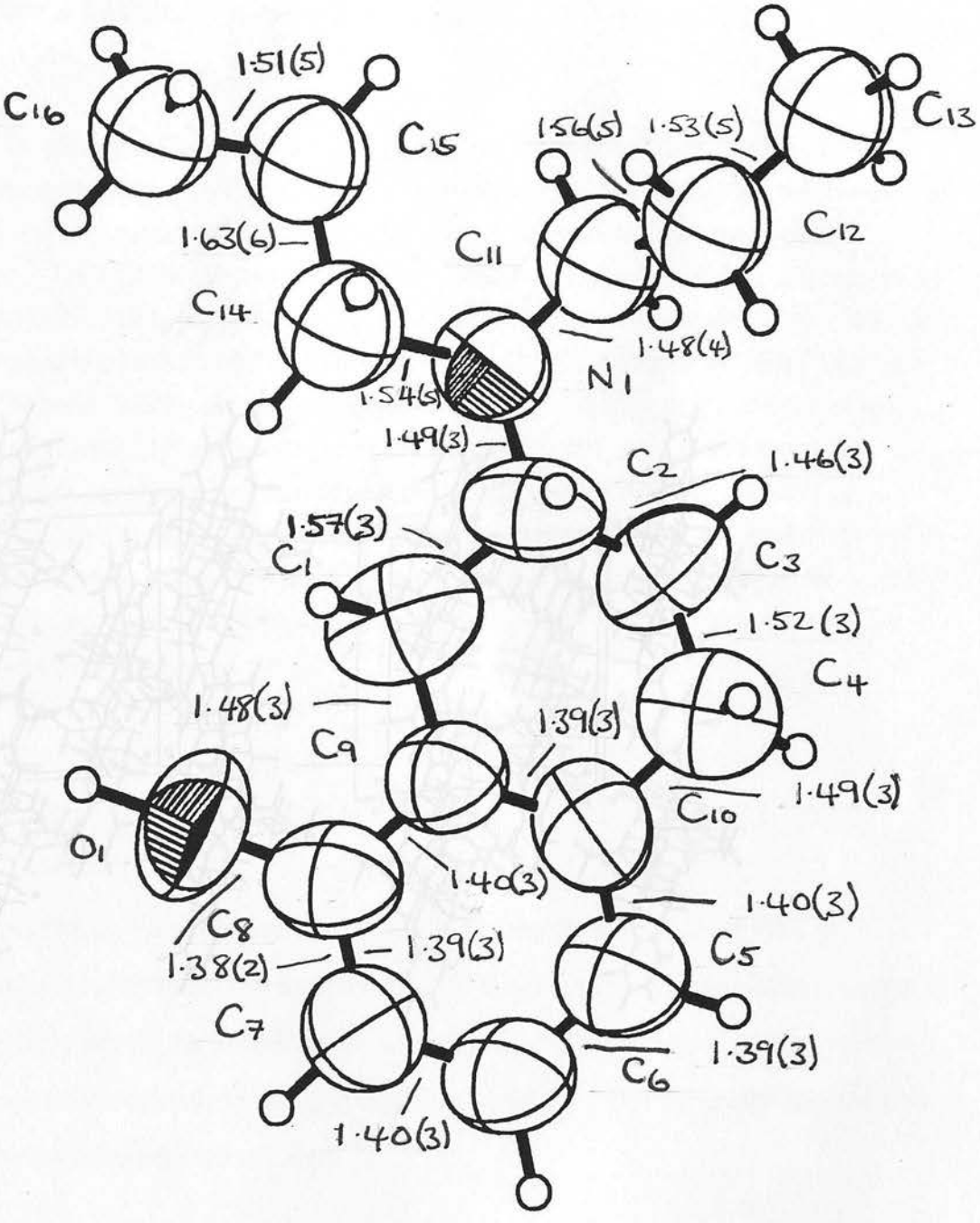


Figure 2.7 Perspective drawing of 8-OH DPAT. Non-Hydrogen atoms are shown as 50 % thermal ellipsoids; O and N atoms are shaded.

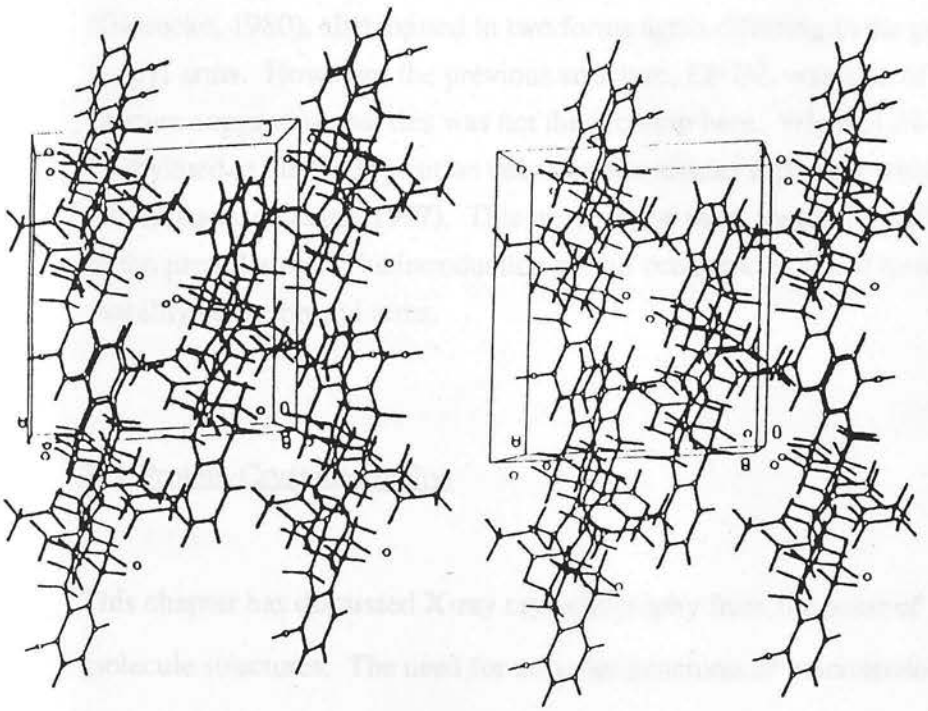


Figure 2.8 Molecular packing of 8-OH DPAT, c-axis projection.

ring, however only 4 out of 6 atoms on the non-aromatic ring superimpose. The propyl arms, C (11) to C (13) lie in the same direction but their atoms do not superimpose. Superimposing along the N (1) to C (11) bond shows a deviation of 20° for the C (12) atom when viewed from the C (11) position.

The disorder of one, or both, of the propyl arms in this structure has hindered the structure solution. Only one crystal was ever obtained and the mixture easily oxidised in solution rendering further attempts unsuccessful. A previous structure 8OH-DPAT.HBr (Karlsson *et al.*, 1988) which existed in two forms, differing in the position of their propyl arms, was of one enantiomer within the asymmetric unit. A structure with the hydroxyl transferred to the C (5) position, 5OH-DPAT.HCl (Giesecke, 1980), also existed in two forms again differing in the position of the propyl arms. However, the previous structure, EP092, was also of a racemic mixture suggesting that this was not the problem here. When 5OH-DPAT is methylated at the C (3) position only one enantiomer is present within the asymmetric unit (Johansson *et al.*, 1987). This methylation must somehow stabilise the mobility of the propyl arms. The introduction of half occupancies failed to account for the mobility of the propyl arms.

2.5 Protein Crystallography.

This chapter has discussed X-ray crystallography from the point of view of small molecule structures. The need for accurate structures of macromolecules is inevitable in the desire for the knowledge of how biological systems function. The transition from small molecule to macromolecular structures is relatively straightforward, but a few modifications are required.

The growth of protein crystals is often the rate limiting step in the proceedings. Crystals tend to be fragile and have a short beam lifetime. For this reason very intense synchrotron X-ray sources are often used to allow quicker data collection.

Synchrotrons allow data collection to a higher resolution and can be tuned such that the wavelength is between 0.5 Å and 2.0 Å approximately.

The phase problem can be overcome in much the same way as before. Generally the introduction of heavy atoms is used to give initial phase estimates by multiple isomorphous replacement (Chapter 6 in Blundell & Johnson, 1976). Molecular replacement (Rossmann, 1990), the use of a known structure as a trial model and set of phases, is becoming more common as families of proteins begin to appear.

Owing to the vast number of atoms involved and the limited resolution achievable, electron density interpretation tends to be difficult. The knowledge of bond lengths, angles etc. obtained from years of small molecule work helps to build the geometry of some regions of the protein. Also, the knowledge of protein secondary structure motifs helps initial backbone traces from electron density. Refinement, too, is the same least squares procedure but limited resolution means that constraints and restraints are essential. Molecular Dynamics (MD) refinement (Gros *et al.*, 1989) is becoming more common, and allows the crystallographer to search more regions of conformational space for his model structure. This allows for a greater radius of convergence.

CHAPTER 3

Digestion of ß-lactoglobulin.

Chapter 3.

3.1 Introduction.

Although the function of ß-lactoglobulin (Blg) is unclear, carrier properties have been observed. It is known to bind tightly to a variety of small hydrophobic ligands (Hambling *et al.*., 1991) and it has been shown to increase the uptake of retinol in rats (Said *et al.*, 1989). It is likely that the central, and hydrophobic, core is responsible for these properties.

It would be beneficial if this cavity could be engineered to carry specifically an alternative ligand for example. Various drugs can cause problems due to there insolubility, or they may be easily degraded within the stomach. Some can also cause major irritation and damage to the stomach lining (see Chapter 1.5). It would be of advantage, therefore, if these drugs could be carried through the stomach to the gut without eliciting any of the aforementioned problems. Modified Blg may be the carrier molecule to carry out this function.

If Blg were to be used as a carrier molecule it would be required to pass through the stomach intact with the ligand still bound, allowing its absorption in the intestines.

It is known that Blg is acid stable (Aschaffenburg & Drewry, 1967) and it has been shown that it can survive passage into the intestines (Miranda & Pelissier, 1983). Blg has also been shown to cross successfully rabbit ileum *in vitro* in both intact and degraded forms (Marcon-Genty *et al.*, 1989).

Resistance to stomach proteases and the effect of a bound ligand on this resistance to proteolysis is investigated in this chapter over a time period concomitant with its incubation within the stomach.

With a ligand bound, the protein may take up a slightly different conformation. This change may affect the digestion pattern of the protein and must therefore be investigated.

The effect, on digestion, of a ligand bound to a protein has been previously studied with creatine kinase (Jacobs & Cunningham, 1968) and with cholera toxin (Galloway & van Heyningen, 1987). In both cases, different digestion patterns were shown. Here, retinol was chosen as a suitable ligand for studies on the digestion of Blg in the presence and absence of ligand because of its affinity for the protein (Fugate & Song, 1980). Of all the ligands known to bind to the protein, it has the highest affinity constant (Hambling *et al.*, 1991).

In most cases the digestion is undertaken at the optimum pH for the enzyme being used. Some peptic digests were performed at pH values higher than basal stomach levels as Miranda & Pelissier (1983) indicate that the pH is raised on the ingestion of food. Fresh solutions of Blg were prepared for each investigation since storage at some pH's can cause irreversible denaturation (Akroyd, 1965).

The investigation is dependent upon the enzymatic cleavage of the polypeptide at a limited number of sites such that a few small peptides are generated. The gel system of Laemmli (1970) gave inadequate resolution of peptide bands. For this reason the gel system of Schagger & von Jagow (1987) was followed. This, in conjunction with the minigel system recommended by Plaxton & Moorhead (1989), allowed a resolution of smaller molecular weight peptides at lower acrylamide concentrations.

3.2 Materials and Methods.

Bovine ß-lactoglobulin (Blg) was obtained from either SIGMA or Pentex, both of which comprised a mixture of the A and B variants. Digestion buffers used were 50mM citrate/Na_2HPO_4 at the appropriate pH. In all cases proteolysis was carried out at 37 °C in a shaking water bath. Samples were diluted 1:2 with sample buffer (4% SDS, 12% glycerol w/v, 50mM TRIS, 2% mercaptoethanol v/v, 0.01% bromophenol blue, adjusted to pH 6.80 with HCl) before being run on an SDS-polyacrylamide gel.* Different gel compositions were tried using the method of Schagger & von Jagow (1987) with 16.5% T, 3% C gels giving the best resolution of peptide bands. T denotes total percentage concentration of both monomers (acrylamide and bis-acrylamide, both from Bio-rad) and C denotes the percentage of cross linker relative to the total concentration. Gels were stained in 0.25% Coomassie blue in 50% methanol, 10% acetic acid for 1 hour before washing in 5% ethanol, 7% acetic acid until the background became clear.

3.2.1 Pepsin Digestion.

Blg (2mg/ml) was dissolved in digestion buffer at pH 2.0 (pH lowered by acetic acid). Bovine pepsin (SIGMA) was dissolved in digestion buffer at pH 2.0 (0.5mg/ml) and frozen in aliquots. Reactions were initiated by addition of 50µl pepsin to 1ml of the Blg solution giving a 300:1 molar ratio of Blg to pepsin. Samples of 100µl were removed at intervals and the reaction stopped by addition of concentrated KOH to raise the pH of the reaction medium to well above that at which pepsin is active. Control experiments under identical conditions used 2mg/ml α-lactalbumin (α-la) at a 390:1 molar ratio of α-la to pepsin.

* Molecular weight standards used were from Pharmacia, and were as follows: Phoshphorylase b (94,000 D), Albumin (67,000 D), Ovalbumin (43,000 D), Carbonic anhydrase (30,000 D), Trypsin inhibitor (20,000 D), α-lactalbumin (14,400 D).

3.2.2 Trypsin digestion.

Blg was dissolved in digestion buffer at pH 8.0 (2mg/ml). Bovine pancreatic trypsin (SIGMA T0134) was dissolved in pH 8.0 digestion buffer (0.5mg/ml) and frozen in aliquots. Reactions were initiated by addition of 20µl trypsin to 1ml of the Blg solution giving a 260:1 Blg to trypsin molar ratio. Samples of 100µl were removed at intervals and the reaction stopped by addition of PMSF (phenylmethanesulphonyl fluoride) in propanol to a final concentration of 10mM. Control experiments under identical conditions used α-la (2mg/ml) at a 340:1 molar ratio of α-la to trypsin

3.2.3 Gastric Fluid Digestion.

Human gastric fluid was obtained from the Western General Hospital, Edinburgh by Dr Andrew Ryle. Blg was dissolved at 2mg/ml in digestion buffer at pH 3.0 and 40µl of gastric fluid added to 1ml of reaction mixture. Samples of 100µl were removed at intervals and the reaction stopped by addition of concentrated KOH to raise the pH of the reaction medium. The amount of gastric fluid to add was estimated using a 1 mg/ml solution of α-la under identical conditions, which also served as a control.

3.2.4 Ligand Bound Experiments.

Proteolysis was carried out in a manner similar to that previously described for trypsin digestion. Digestion buffer used was pH 7.5 and protease concentration was half of that originally stated to slow the effect of trypsin and allow any effect of the retinol to be seen. Digestions contained retinol (20, 60 and 100µl) from a 20mM

stock in ethanol. This allowed reasonable ratios of retinol to Blg, (of 4:1, 22:1 and 37:1 respectively). Hemley *et al.*, (1979) recommend that Blg should not be used in solutions of ethanol that exceed 5% as the protein becomes denatured. For this reason the final percentage of ethanol in solution was never greater than 5%. Digests were repeated with ethanol alone. A preincubation of protein with retinol was allowed for 30 minutes prior to the initiation of the reaction. Absorbances of the protein-retinol mixtures were measured at 350nm on a Pye Unicam SP6-500 UV spectrophotometer before and after the addition of the retinol in order to ascertain if the Blg was removing retinol from solution.

3.3 Results.

3.3.1 Pepsin Digestion.

Preliminary incubations of a Blg solution (2mg/ml) with varied amounts of bovine pepsin (10 µl increments up to 100 µl of a 0.5mg/ml solution) showed no significant difference in digestion pattern of the protein . Digestions therefore used 50µl of pepsin over a period of 3 hours and samples run on a polyacrylamide gel indicated no significant depletion of the major 18kD band after this time period. Figure 3.1 shows digestion up to 3 hours and indicates a lower molecular weight band, around 15kD, on protease addition in all lanes. Presumably this is due to a nicking of some protein molecules leading to a slight reduction in the molecular mass. A solution of α-la at 2mg/ml treated under identical conditions was digested to such an extent that after 1 hour no bands were evident on the gel (Figure 3.2). Blg was dissolved at pH 8.0 at 2mg/ml in digestion buffer, which was then reduced to pH 2.0 using acetic acid. Reaction with 50µl pepsin solution showed no significant difference in the pattern obtained.

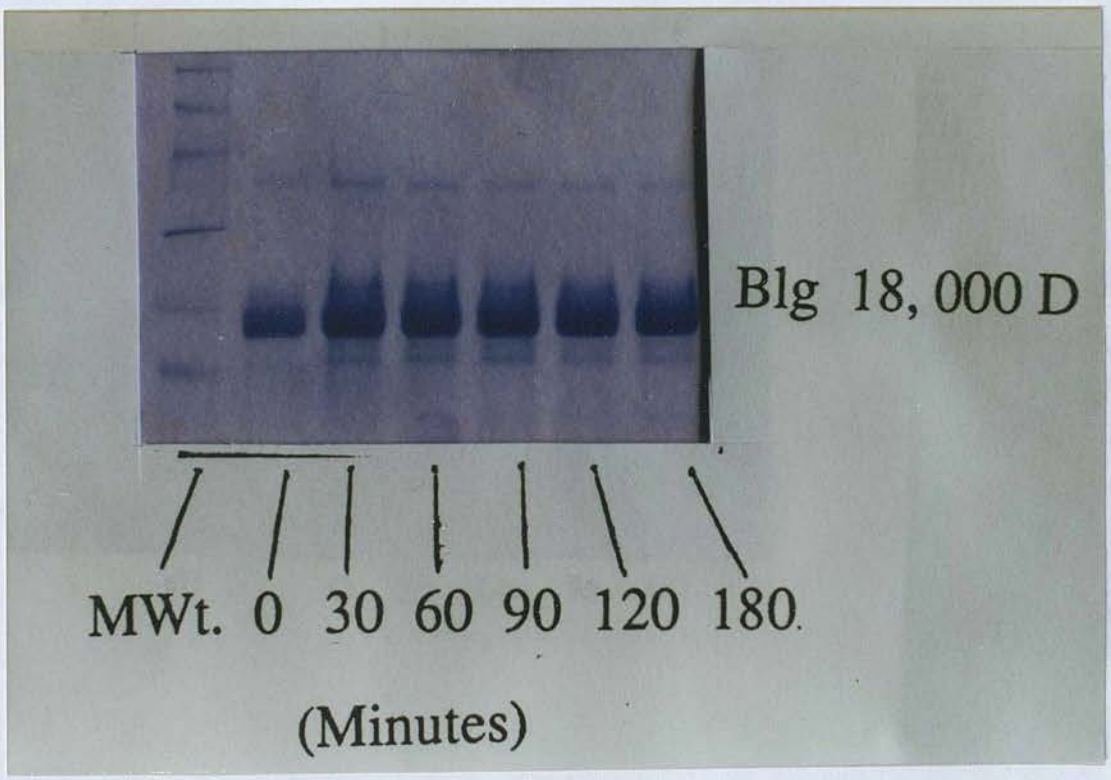


Figure 3.1 Bovine pepsin digestion of bovine Blg at 37°C, pH 2.0 over 3 hours. Molar ratio of Blg to pepsin was 300:1.

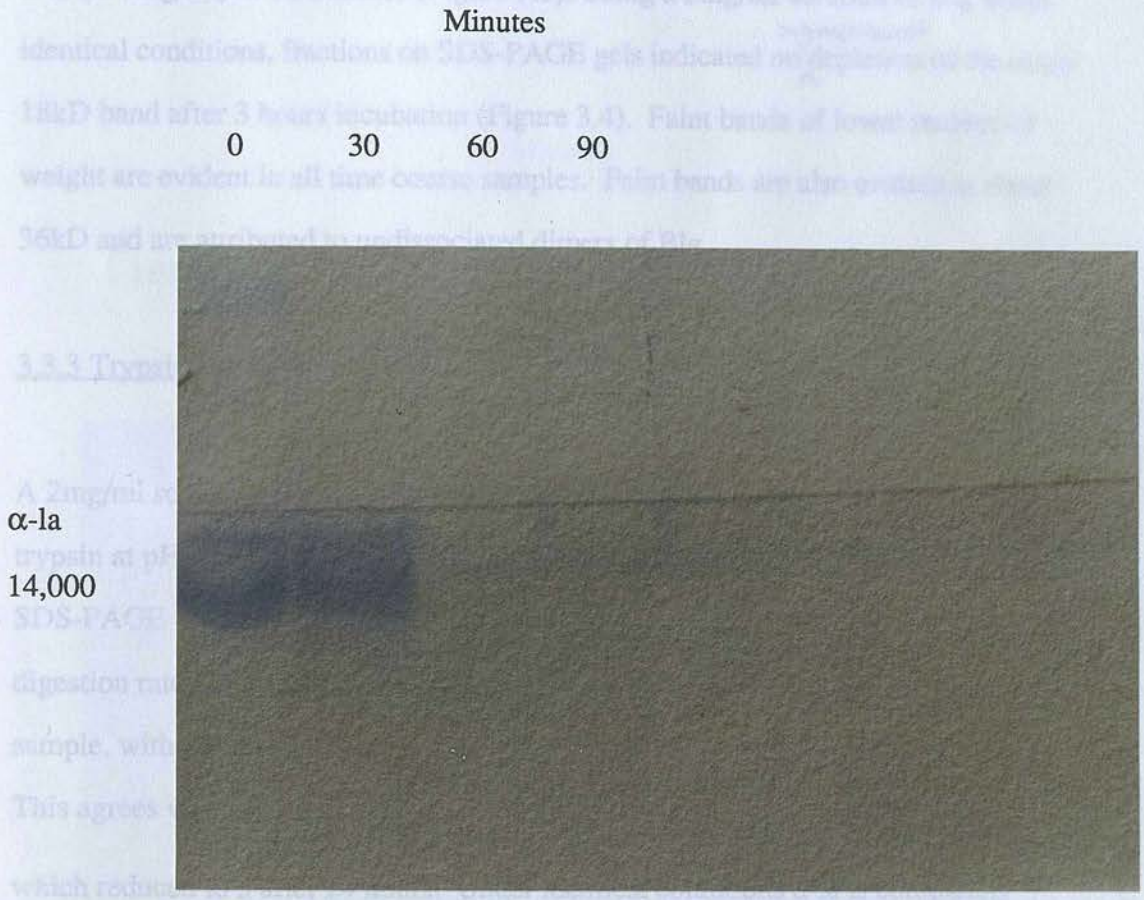


Figure 3.2 Bovine pepsin digestion of bovine α-la at 37°C, pH 2.0 over 2 hours. Molar ratio of α-la to pepsin was 390:1.

3.3.2 Gastric Digestion.

The amount of gastric fluid required for digestion was determined empirically using α-la. 50µl of gastric fluid per ml digestion mix at pH 3.0 gave complete digestion of α-la (at 1mg/ml) in 30 minutes (Figure 3.3). Using a 2mg/ml solution of Blg under identical conditions, fractions on SDS-PAGE gels indicated no significant depletion of the major 18kD band after 3 hours incubation (Figure 3.4). Faint bands of lower molecular weight are evident in all time course samples. Faint bands are also evident at about 36kD and are attributed to undissociated dimers of Blg.

3.3.3 Trypsin digestion.

A 2mg/ml solution of Blg showed some resistance to digestion with bovine pancreatic trypsin at pH 8.0. Figure 3.5 shows the presence of the major 18kD band on SDS-PAGE up to 2 hours, indicating that even though the intensity is decreasing the digestion rate is not rapid. After 30 minutes incubation, 4 bands are obvious in the sample, with a reduction in intensity of the major ones from the start of the reaction. This agrees with Huang *et al,* (1985) who found 4 bands after 4 hours incubation, which reduced to 3 after 24 hours. Under identical conditions α-la is completely degraded after 25 minutes, showing no presence of a major protein band (see ligand bound result, 3.3.4).

3.3.4 Ligand Bound Results.

Trypsin digestion was undertaken at pH 7.5 to allow the effect of the retinol to be seen. The solution of Blg appeared less opaque than the α-la/retinol solution and the

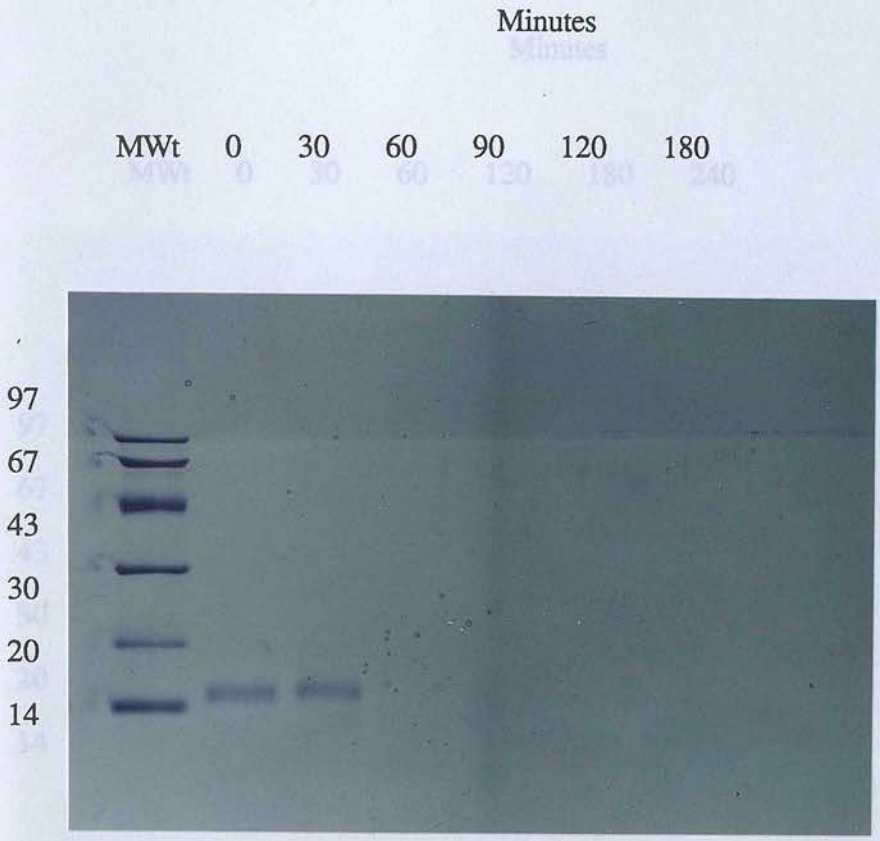


Figure 3.3 Human gastric fluid digestion of bovine α-la (mg/ml) at 37°C, pH 3.0. MWt standards are x 1000 D.

Figure 3.4 Human gastric fluid digestion of bovine Blg (2mg/ml) at 37°C, pH 3.0 over 4 hours. MWt standards are x 1000 D.

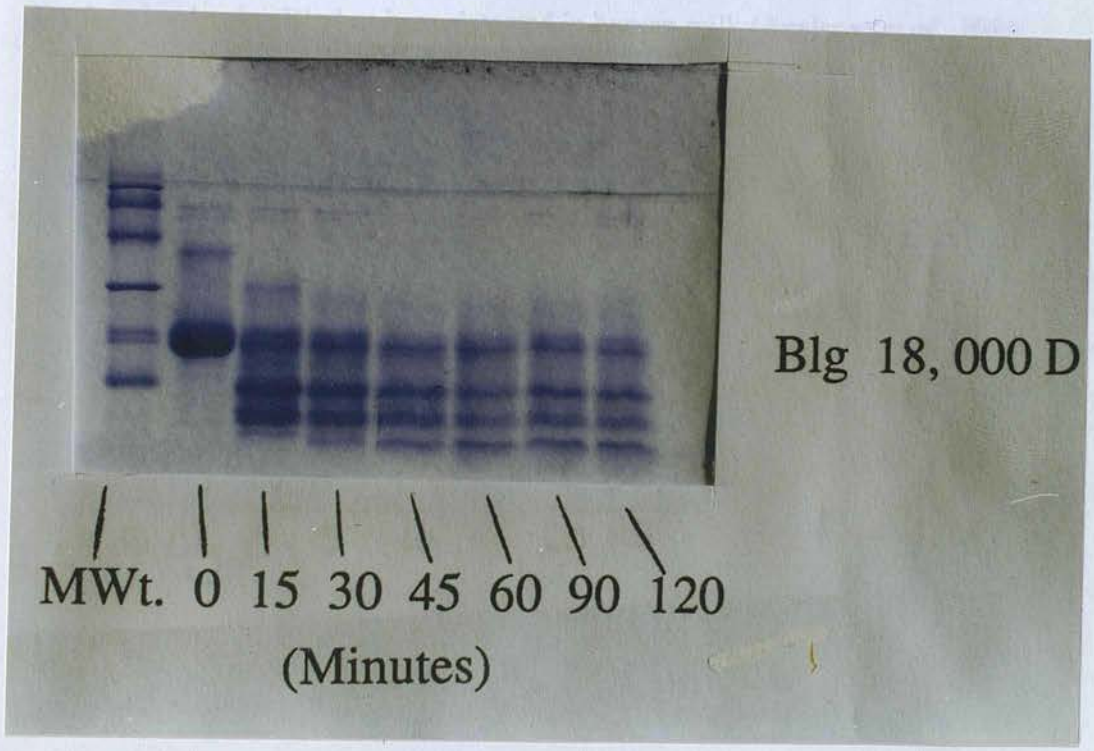


Figure 3.5 Bovine trypsin digestion of bovine Blg at 37°C, pH 8.0 over 2 hours. Molar ratio of Blg to trypsin was 260:1.

absorbance at 350nm for Blg was 23% less than that for the α-la. After digestion for one hour, protection was only evident at the two higher ratios (Figure 3.6). The trypsin digest with ethanol alone showed slightly quicker digestion of native Blg than a digest in buffer alone. This is presumably a result of slight unfolding induced by the ethanol. Under identical conditions, using the same molar ratio of α-la to retinol, digestion was complete after 25 minutes showing that the presence of retinol had no effect in this case (Figure 3.7).

3.4 Discussion.

The fact that bovine Blg has been detected in human milk (Axelsson *et al.*, 1986; Jacobsson *et al.*, 1985; Brignon *et al.*, 1985) and in the urine of new born calves (Pierce, 1960) suggests that Blg crosses the gut wall in some immunologically identifiable form. However, it was not until Marcon-genty *et al.*,(1989) were investigating food protein antigens that it was realised that 6 to 9% of the Blg that traverses the epithelial membrane was in the intact form.

Resistance to low pH is borne out in the purification procedure (Aschaffenburg & Drewry, 1967) of the protein and in this study we found resistance of the protein not only to bovine pepsin but to human pepsins in the form of gastric fluid.

Resistance to proteases at such a low pH must be as a result of the proteins conformation in solution. Below pH 3.5, Blg is known to dissociate into monomers (McKenzie, 1971). The increased net positive charge is thought to result in H-bonding due to titration of carboxyl groups. Kella and Kinsella (1988) believe that this accounts for the remarkable stability at low pH. Proteolytic targets for pepsin are residues such as tryptophan, tyrosine, phenylalanine, leucine and isoleucine. The

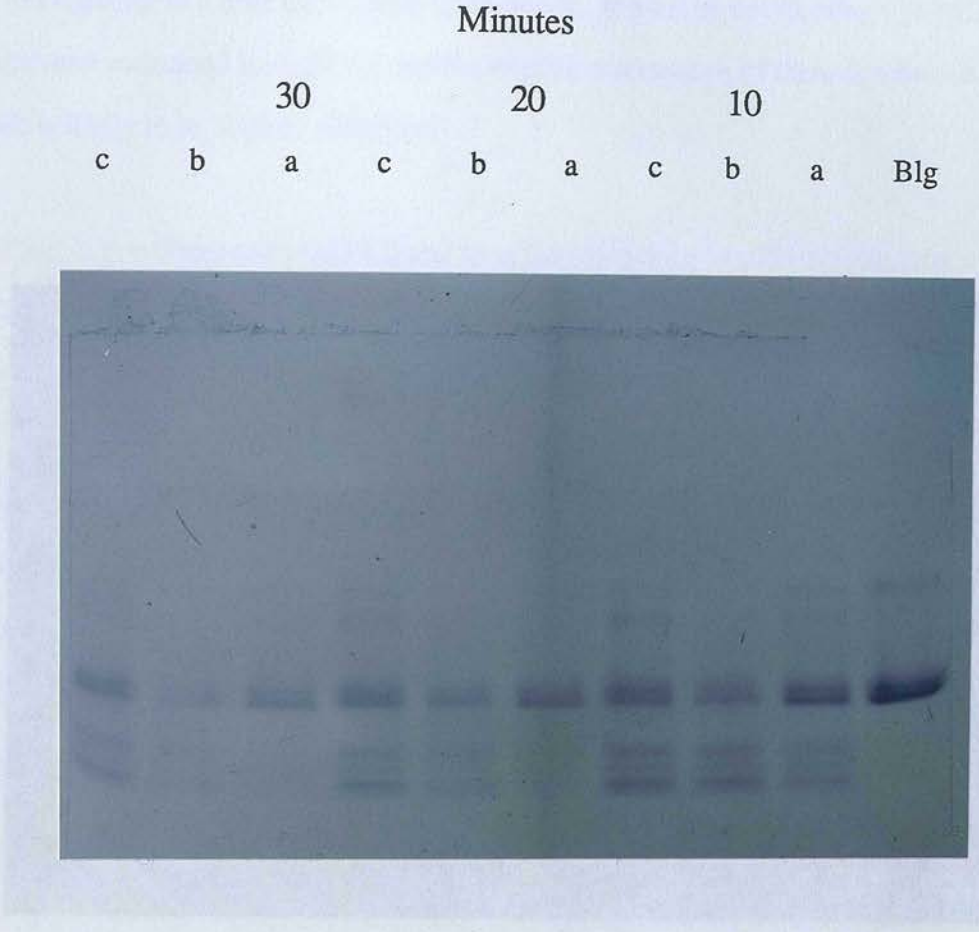


Figure 3.6 Bovine trypsin digestion of bovine Blg (+ retinol) at 37°C, pH 7.5 over 1 hour. Molar ratio of retinol to Blg was 37:1.
a = Blg only; b = Blg + ethanol; c = Blg + ethanol + retinol.

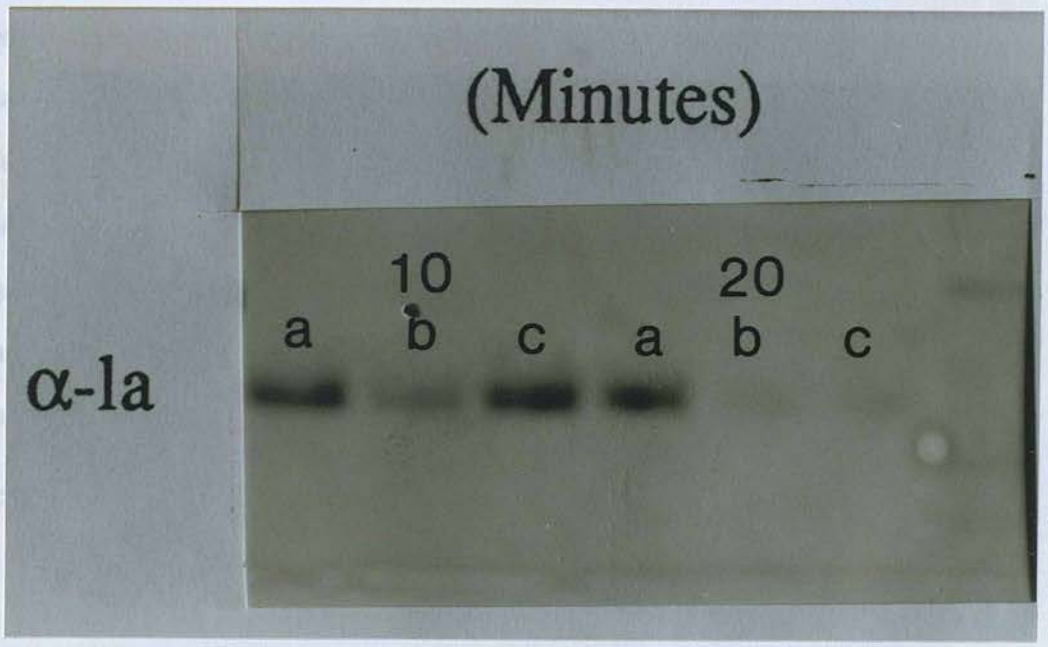


Figure 3.7 Bovine trypsin digestion of bovine α-la (+ retinol) at 37°C, pH 7.5 over 20 minutes. Molar ratio of retinol to α-la was 37:1. a = α-la only; b = α-la + ethanol; c = α-la + ethanol + retinol.

strong resistance suggests that these residues are inaccessible to the enzyme molecules. Examination of the positions of these residues on the lattice Y structure (Figure 3.8) indicates that many of these residues are in fact exposed. The appearance of some lower molecular weight peptides (Figure 3.1) suggests that some peptides are released. If many more sites are attacked then this suggests that they are held together in a near native state by H-bonds. It must be noted, however, that the structure examined is at pH 7.6 and the relative orientations of these residues at low pH is likely to be slightly different.

When Blg is dissolved at pH 8.0 and then the pH reduced to 2.0, for reaction with pepsin, no change in the susceptibility of the protein is seen. This indicates that any conformational changes that occur over this pH range are reversible and in no way destabilise the protein. It is possible that at pH values greater than pH 9.0 would have had a different effect since irreversible changes occur owing to the ionisation of tyrosine residues (Townend *et al.*, 1969). Various biologically active milk proteins, such as immunoglobulins (Brown *et al.*, 1976), transferrin (Brock *et al.*, 1976), lactoferrin and lysozyme (Kato *et al.*, 1985) have previously been shown to be resistant to proteolytic digestion. This is not seen here with α-la, which was used as a control since it is of similar size and origin to Blg.

If Blg is to be considered for use as a drug carrier, the effect of having a ligand bound must also be investigated. Since resistance to trypsin was shown to be less than to the peptic proteases, the effect of a bound ligand on digestion by trypsin may be more marked.

Trypsin is seen to reduce the protein to three smaller molecular weight peptide bands, which gradually decrease in intensity over time. The target residues, lysines and

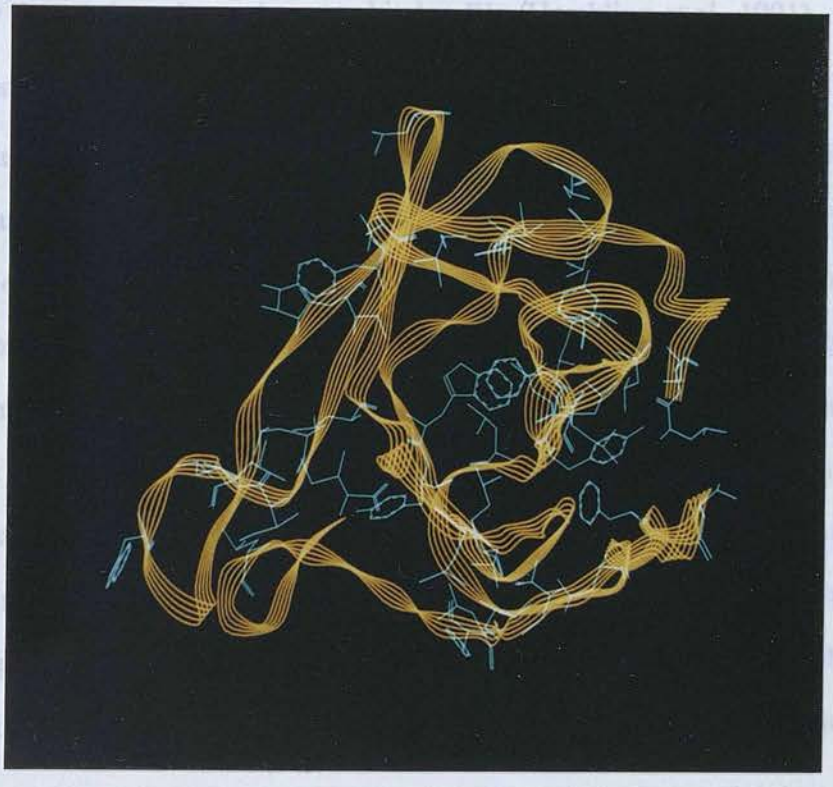


Figure 3.8 Positions of the sites of pepsin attack on the bovine Blg model.

arginines, appear on the lattice Y model structure (Papiz *et al.*, 1987) on the surface down two sides of the molecule, so it is not entirely clear why the protein still shows some form of resistance when compared to α-la under identical conditions.

McKenzie and Sawyer (1967) indicate that a single group becomes ionised in both A and B variants between pH 6.5 and 9.0, because of the release of an anomalous carboxyl group, and at higher pH's the dimer tends to dissociate to monomer. This structural change may account for the greater susceptibility to trypsin relative to pepsin. This is not entirely clear, however, since the loss of hydrogen bonds will result in their replacement by ion pairs which are stronger.

Many ligands have been shown to bind to Blg (Hambling *et al.*, 1991). Retinol was chosen for this study as a result of its high affinity for the protein ($K_a = 5 \times 10^7\ M^{-1}$, Fugate & Song, 1980). The pH of hydrolysis was lowered to pH 7.5, and the protease concentration decreased to reduce the activity and allow the reaction to be more easily monitored. The effect of pH on hydrolysis by trypsin has been monitored (Chobert *et al.*, 1991) and a preference for some sites over others was indicated in the pH range 7.5 to 9.0. The results here indicate that at ligand to protein ratios of 22:1 and 37:1 protection effects are seen. Protection is shown by the greater intensity of bands with retinol present when compared to the digestion pattern with trypsin alone (Figure 3.6). Indeed the samples with ethanol show greater susceptibility to trypsin, with fainter bands presumably indicating an unfolding of the protein. The effect of retinol is shown to be one of protection, and not inhibition of the trypsin, by using identical conditions with α-la, where no effect is seen.

Interactions with the retinol allow the protein structure to be held together, even on hydrolysis of some surface peptides. These interactions may even cause a tightening

of the globular structure allowing the protection of normally susceptible peptide bonds.

The trypsin experiments are not relevant to the passage of the Blg through the stomach. Blg will only encounter trypsin within the duodenum. They do, however, show the effect of a ligand bound to the protein in the process of digestion.

Once the protein bound-ligand reaches the intestines, it has passed a major danger area and the ligand would hopefully be absorbed there. The ligand may even be internalised along with the protein.

3.5 Conclusions and Future Work.

From this study it can be concluded Blg shows remarkable resistance to low pH and to the effects of acidic proteases. Its resistance to trypsin proteolysis is enhanced by the presence of a bound ligand. All of these observed properties suggest that Blg is a good candidate to engineer and ultimately use as a carrier protein in an acid media.

The elucidation of a low pH crystal form will aid in the investigation of the proteins properties at low pH and may even give an indication as to why the protein shows such marked resistance.

CHAPTER 4

Antibody Purification.

Chapter 4

4.1 Introduction.

4.1.1 The Immune System.

Mammals, in order to survive, have evolved a defence mechanism. Their first mode of defence is a barrier, i.e. the skin. Their next line of defence is somewhat more complex and can be split into two parts. Humoral immunity is conferred by cells which, when encountering a pathogen, secrete a molecule known as immunoglobulin and cellular immunity involves intact body cells only, i.e. no freely soluble molecules take part.

Within the humoral system many cells circulate the body containing the information which allows them to bind to any antigen encountered. The antigen, perhaps a foreign protein, becomes bound by a surface receptor on a B-lymphocyte. A region on the surface of the antigen stimulates the B-cells to multiply and produce many clones. Each of the unique B-cells secrete an active form of their surface receptor, the immunoglobulin (Ig). Each B-cell produces a unique Ig which binds to the antigen by its complimentary determining region (CDR). This site on the antigen that stimulates the B-cell need not be antigenic, i.e. it need not be the site that ultimately interacts with the antibody. Some antigenic sites are silent immunogens (Lerner, 1982).

Dependent on the type of response different classes of immunoglobulin are produced each of which has a slightly different conformation. The immunoglobulins are a class of tetrameric proteins with 2 heavy and 2 light chains joined by disulphide bonds. Each chain has two regions. The variable region contains the antigen

combining site and the constant region is invariant within a given class or subclass and is responsible for effector functions (Figure 4.1).

The genetics of the immunoglobulins is complex as it requires the splicing together of a large number of exons to obtain one antibody molecule. Gene shuffling and splicing ensure a different amino acid sequence for each specific antibody. How an antigen stimulates the correct sequence of events is not fully understood.

Present understanding of antigen/antibody interactions comes from X-ray studies of Fab-antigen complexes. Structures have been solved for complexes with hen egg white lysozyme (Amit *et al.*, 1986; Sheriff *et al.*, 1987) and the viral antigen influenza neuraminidase (Colman *et al.*, 1987).

The Fab complex structures with lysozyme (Amit *et al.*, 1986; Sheriff *et al.*, 1987) suggest a 'lock and key' mechanism but that of the influenza neuraminidase (Colman *et al.*, 1987) resembles a handshake,with a sliding motion of the antibody domains. An interaction area of 20 x 30 Å is seen between antibody and antigen in the lysozyme complex but is undefined by Colman *et al.* (1987) for the neuraminidase complex. Modelling of the third CDR of both heavy and light chains in the neuraminidase antibody should resolve further interactions in this structure. Definite antigen distortion is apparent, however, with Cα atoms being moved from their native N9 neuraminidase positions by 1Å or more. Further data should establish the importance of V_L and V_H sliding in immune recognition. (Figure 4.1)

The clinical interest in bovine milk is a result of the severe allergic reactions it can produce in infants. These reactions have been attributed to the Blg component within the milk. As a result it should not present a problem in producing antibodies

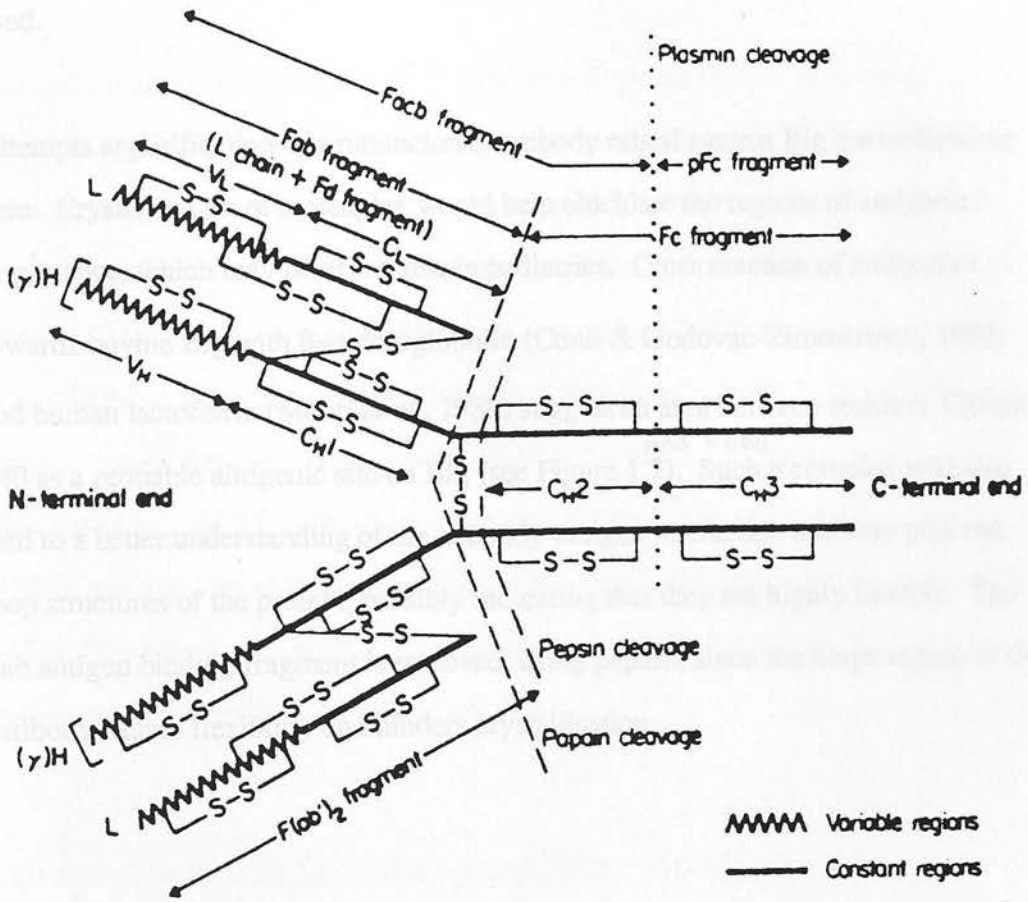


Figure 4.1 Line drawing of immunoglobulin G indicating the different structural domains.

towards the protein. In this Chapter the production of a polyclonal antiserum (i.e. Ig molecules from a variety of B-cell clones) is described. The polyclonal antibodies are purified and used to show the validity of the techniques being used. Crystallisation of an antibody from polyclonal antisera is difficult due to the many species of antibodies that exist therein. For this reason a monoclonal antibody source, a unique antibody variant synthesised by the classic Kohler and Milstein technique (1975), is used.

Attempts at purification of a monoclonal antibody raised against Blg are undertaken here. Crystallisation of a complex would help elucidate the regions of antigenic importance, which may be of interest in pediatrics. Cross reaction of antibodies towards bovine Blg with β_2-microglobulin (Conti & Godovac-Zimmerman, 1990) and human lactoferrin (Monti *et al.*, 1989) suggest an area between residues 120 and 140 as a probable antigenic site on Blg (see Figure 1.2 and 7.2(b)). Such a complex will also lead to a better understanding of the antibody-antigen interaction and may pick out loop structures of the protein, possibly indicating that they are highly flexible. The Fab antigen binding fragment is removed, using papain, since the hinge region of the antibody causes flexibility and hinders crystallisation.

4.2 Materials and Methods.

4.2.1 Purification of a Monoclonal Antibody from Mouse Ascites Fluid.

Mouse ascites fluid (1ml) was dialysed against running buffer (20mM Tris-HCl, pH 7.5) before loading onto a DEAE Affi-gel Blue column (Bruck *et al.*, 1982). The ascites fluid was loaded onto a 5ml column using three bed volumes of running

buffer and 50 fractions of 2ml read for absorbance at 280nm. Globular proteins were then eluted either stepwise, with 25, 50 and 75mM NaCl in running buffer, or by using a gradient of 0-100mM NaCl in running buffer. Fractions were analysed by SDS PAGE or ELISA.

4.2.2 Fab Preparation.

Isolated immunoglobulin (IgG) antibodies were dialysed against three changes of digestion buffer (500ml, 0.1M sodium phosphate, 0.01M cysteine, 2mM EDTA, pH 7.0) over a period of 24 hours. Immobilised papain (2mg) was suspended in 1ml digestion buffer and the reaction initiated by addition of an equal volume to the IgG mixture. The reaction was allowed to continue at room temperature for 4 hours before being stopped by spinning down the immobilised papain. The immobilised enzyme could be reused. The Fab was purified by HPLC, by applying the supernatant to a DEAE-cellulose column, equilibrated with 5mM Sodium Phosphate, pH 8.0, and was eluted using a gradient of 5mM to 0.3M Sodium Phosphate buffer, pH8.0. The absorbance at 280nm of eluted fractions (2ml) was read and these were further examined by SDS-PAGE and ELISA.

4.2.3 Immobilising β-lactoglobulin for an Affinity Column.

A bovine ß-lactoglobulin (Blg, SIGMA) affinity column was made using Cyanogen bromide (CNBr) activated Sepharose (SIGMA) as a support matrix. CNBr sepharose was activated by suspending 1g of the sepharose in 200ml of 10mM HCl on a sintered glass funnel for 15 minutes. Coupling of Blg to the matrix was achieved by addition of 1g activated sepharose to 5ml of a 2mg/ml solution of Blg in 0.1M $NaHCO_3$, 0.5M NaCl. The solution was mixed for 1 hour. The percentage of

protein bound was estimated by measuring the absorbance at 280nm using an extinction coefficient of 0.96 M^{-1} cm^{-1} (Townend *et al.*, 1960). Unreacted groups on the Sepharose were blocked by mixing with 0.1M Tris-HCl (or 1.0M ethanolamine), pH 8.0 for 1 hour. Three washing cycles were then employed to remove non-covalently linked adsorbed protein, with each cycle containing a wash at pH 4.0 (0.1M acetate buffer, 0.5M NaCl) followed by a wash at pH 8.0 (0.1M Tris, 0.5M NaCl).

4.2.4 Raising a Polyclonal Antiserum.

All parts of the following procedures involving the animal were performed by Dr Gordon Harkiss from the Department of Veterinary Pathology, The Royal (Dick) Veterinary School, Edinburgh.

Bovine Blg was added to Freunds complete adjuvant (100mg/ml) and mixed to form an emulsion. The emulsion was then injected intramuscularly into a rabbit with booster injections being given 10, 20 and 30 days after the initial injection. The animal was bled, initially after 20 days, from a vein in the ear and the blood was left to stand at 4 °C. Once the red cells had clotted the serum was removed and centrifuged at high speed to remove further cellular material. The serum was then tested for activity towards Blg using an ELISA.

4.2.5 Enzyme-linked Immmunosorbant Assay (ELISA).

To screen for the presence, or activity, of an antibody a plate binding ELISAwas used. Microtitre plates (Flow laboratories) were coated overnight at 4 °C with

75µl/well of antigen in coating buffer (15mM carbonate buffer, pH 9.6), washedwith phosphate buffered saline (140mM NaCl, 2mM NaH_2PO_4, 7mM Na_2HPO_4, pH 7.2-7.4; PBS)/0.01% Tween 20, before being blocked with 1% gelatin in PBS. Primary antibody was added (75µl/well) and incubated for 2 hours at room temperature. The plate was washed and 75µl/well of a 1/1000 dilution of a second antibody conjugated to HRP (Horse Radish Peroxidase) or AP (Alkaline Phosphatase) was added. Positive wells were rapidly screened due to a colour change on the addition of the appropriate substrate (see below):

Para nitro-phenyl phosphate (PNPP) buffer. (for AP conjugate).	Ortho-phenylene diamine (OPD) (for HRP conjugate)
8mg of PNPP	8mg of OPD
10ml of 0.5mM $MgCl_2$	2.5ml of 0.1M citric acid
5ml of 10mM triethanolamine, pH 9.5	2.5ml of 0.2M Na_2HPO_4
5ml of H_2O	5ml of H_2O
	4µl of H_2O_2 before use

Absorbance could be read off the plates at 400nm (PNPP) or 492nm (OPD) using an ELISA plate reader (Titretek, Flow laboratories).

4.2.6 SDS-Polyacrylamide Gel Electrophoresis (Laemmli, 1970).

SDS-PAGE was performed using 10 to 20% resolving gels made up from a 30% acrylamide stock (acrylamide:bis ratio 30:1) diluted in resolving gel buffer and polymerised by the addition of ammonium persulphate (APS) and TEMED, both at a

final concentration of 0.01%. Following polymerisation, the stacking gel (4% acrylamide) was prepared in stacking gel buffer and polymerised by addition of APS and TEMED to a final concentration of 0.01%. Samples were diluted 1:2 in sample buffer and boiled for 3 minutes prior to loading. Gels were run in Tris/glycine buffer, under constant current conditions, until the bromophenol marker was 1cm from the bottom of the gel.

Acrylamide stock.

58g Acrylamide
2g NN'-methylenebisacrylamide
Made up in 200ml distilled H2O,
filtered through Whatman No.1 filter
and stored in a dark bottle

Stacking gel buffer.

30.25g Tris base
2.5g SDS
made up to 500ml in H_2O

Resolving gel buffer.

90.75 Tris base
2.5g SDS
made up to 500ml in H_2O
and pH'd to 8.8 with 5M HCl

Running Buffer

6g Tris base
28g Glycine
10ml of 10% SDS
Made up to 1 litre with H2O

Sample Buffer

10ml stacking gel buffer
2g SDS
5ml B-mercaptoethanol
10ml Glycerol
5mg bromophenol blue
Made up to 100ml with H_20

4.2.7 Coomassie Blue Staining of Polyacrylamide Gels.

After electrophoresis, gels were fixed in 50% methanol, 10% acetic acid, for 30 minutes before being stained, for 1 hour in 0.25% Coomassie blue in fixer solution. Gels were destained by washing in a solution of 5% ethanol, 7% acetic acid until the background had cleared.

4.2.8 Silver Staining of Polyacrylamide Gels.

After electrophoresis the gels were fixed in 50% methanol, 10% acetic acid for 30 minutes. The gel was washed in 5% methanol, 7% acetic acid before being transferred into 10% glutaraldehyde for 30 minutes. Gels were then washed for at least 2 hours in 5 litres of H_2O before soaking in 0.1% silver nitrate solution for 30 minutes. Excess silver nitrate was removed by rinsing the gel in H_2O before developing in 100ml of 3% sodium carbonate (Na_2CO_3) containing 50µl formaldehyde. Once bands had appeared the reaction was stopped using 5ml 2.3M citric acid. Gels were cleaned by washing in H_2O.

4.2.9 Electroblotting (Western Blot).

Following electrophoresis, the stacking gel was removed and the resolving gel placed in a tray of blotting electrode buffer (20% methanol, 25mM Tris Base). A piece of nitrocellulose membrane was soaked in the buffer before sandwiching both between filter paper and then between the blotting apparatus (Biorad). Electroblotting was carried out at 120mA for 30 minutes. Successful transfer was monitored by the appearance of rainbow molecular weight markers (SIGMA P. 1677) on the nitrocellulose filter.

4.2.10 Immunodetection of Proteins.

The unreacted sites on the nitrocellulose membrane were blocked by incubation in a 0.15% Tween 20/Phosphate buffered saline (PBS) solution for 1 hour. The membrane was then rinsed in three changes of PBS before being transferred to a plastic bag. Primary antibody was added to the bag and allowed to shake for 1 hour. Following this, the membrane was again washed before being transferred to a fresh bag and a 1/1000 dilution of HRP conjugated second antibody added. The bag was sealed and shaken for 1 hour before washing and substrate addition. Substrate was allowed to develop until the appearance of bands, and the reaction terminated by washing in several changes of H_2O.

Substrate.

10mg Diamino Benzadine

9ml PBS

10ml H_2O_2

1ml of 1% cobalt chloride

4.3 Results and Discussion.

Initial monoclonal antibody purification procedures were performed using ascites fluid containing a non-related antibody. The monoclonals were supplied by Dr G. Harkiss (Dept. of Vet. Pathology, University of Edinburgh).

After dialysis against column-running buffer, the ascites fluid was loaded onto a DEAE-Affigel blue column and fractions collected. The stepwise wash for removal of the globular fraction proved more successful than a gradient elution. Figure 4.2 shows two peaks of protein, one after each salt wash. The peaks are attributed to transferrin and immunoglobulin (Ig) respectively (Bruck *et al.*, 1982). The presence of the IgG in peak 2 was confirmed using an ELISA coated with sheep anti-mouse Fab before sample addition. The ELISA was probed with Rabbit anti-mouse Ig (whole molecule) conjugated to HRP (Figure 4.3). Positive and negative controls were used appropriately . All fractions containing Ig were pooled and concentrated to approximately 1ml in volume using an Amicon ultrafiltration cell. The protein was digested with immobilised papain after dialysis. Digestion was stopped after 4 hours by centrifugation which removed the Sepharose bound protease. Figure 4.4 shows both reduced and unreduced material analysed by SDS-PAGE, indicating that a large fraction of the IgG has successfully been digested. Little contamination is evident as a more prominent band would appear at 50kD in the reduced track owing to the presence of undigested heavy chains. Low molecular weight bands result from contaminating peptides and were removed when the digest material was applied to the DEAE cellulose column . Figure 4.5 shows two peaks when the absorbances of the fractions are read at 280nm. An ELISA coated with the fractions (Figure 4.6), and probed separately with anti-mouse Fc and anti-mouse Fab, confirms the first peak to be purified Fab and the Fc region being eluted at the end of the sodium phosphate gradient. An overlap region in the middle of the graph is thought due to undigested Ig G.

Monoclonal antibodies (IgG2B and IgG3) towards Blg were supplied by Dr J. Gavalchin (State Univ. of NY, Syracuse, USA) in ascites fluid. The following procedures were carried out on both classes of monoclonal antibody. Purification

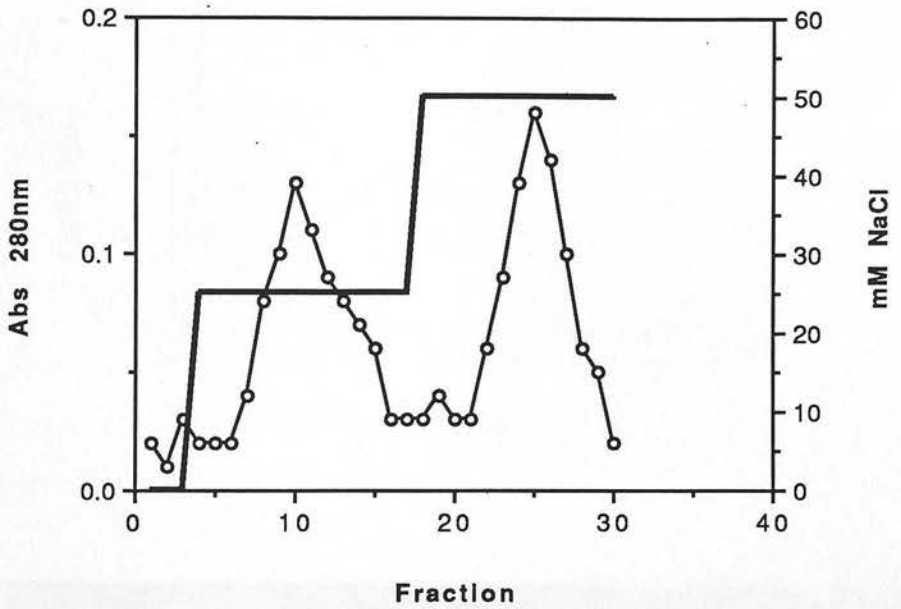


Figure 4.2 Elution profile of mouse ascites from a DEAE - Affi gel blue column using a stepwise elution with NaCl.

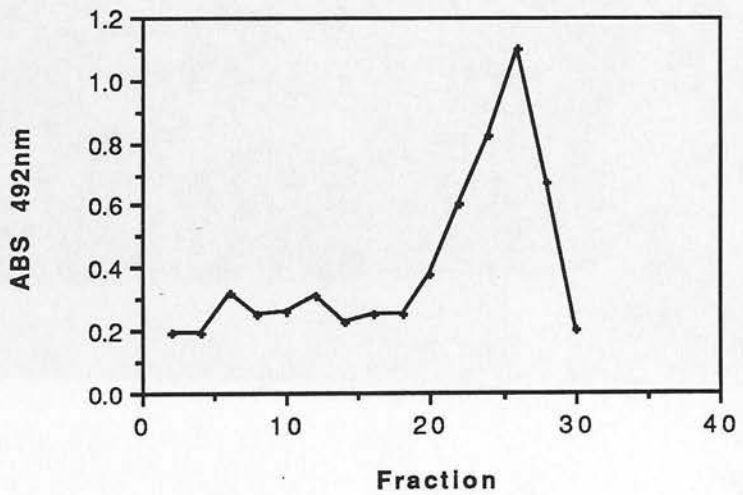


Figure 4.3 ELISA with anti-mouse Ig-G of fractions eluted from DEAE - Affi gel blue column.

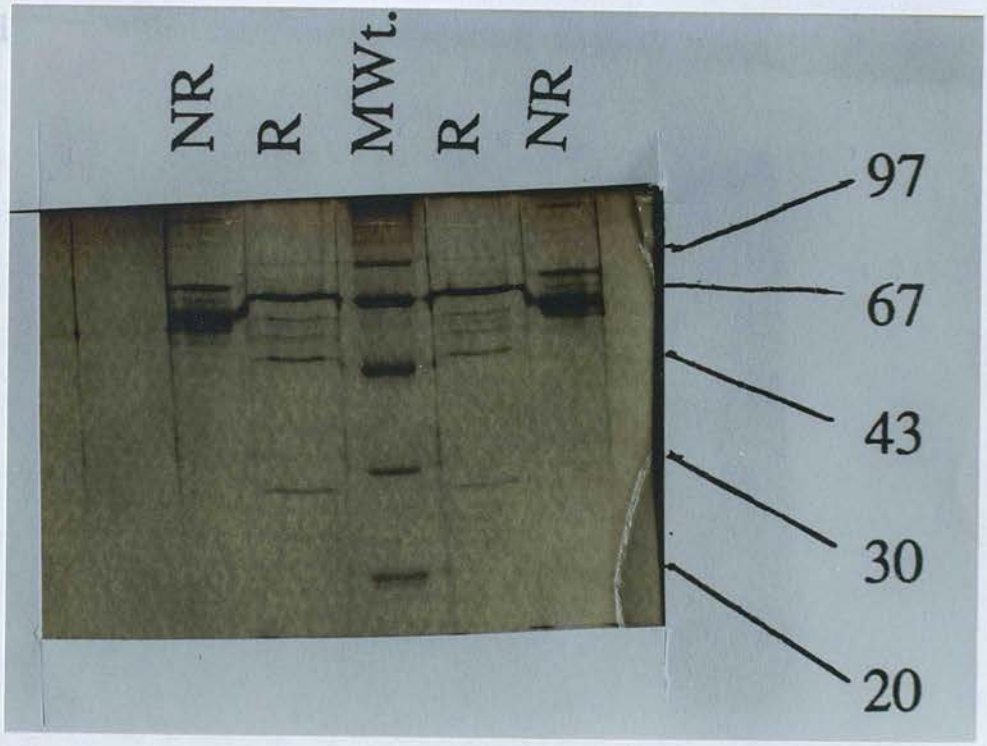


Figure 4.4 Gel of papain digested immunoglobulin in both reduced (R) and non-reduced (NR) forms. MWt = molecular weight markers.

MWt standards are x 1000 D.

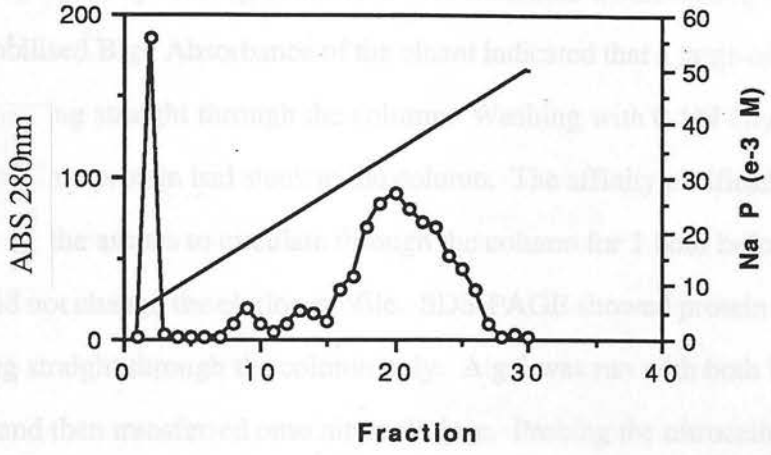


Figure 4.5 Elution profile of digested mouse antibody from a DEAE-cellulose column using a 5 - 50 mM sodium phosphate gradient.

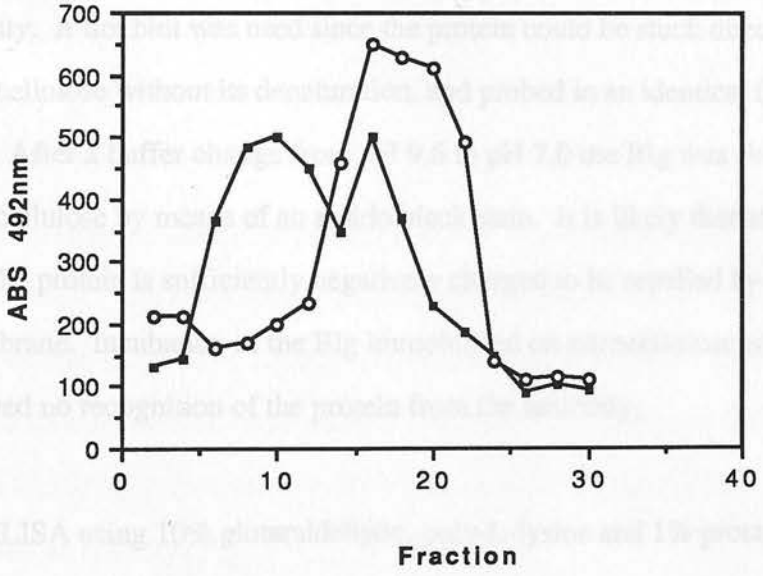


Figure 4.6 ELISA using anti-mouse Fab (■) and anti-mouse Fc (○) against the fractions from the DEAE-cellulose column.

was attempted by loading 100ul of ascites fluid onto a 5ml affinity column with immobilised Blg. Absorbance of the eluant indicated that a large amount of protein was passing straight through the column. Washing with 0.1M Glycine-HCl, pH 2.0 indicated no protein had stuck to the column. The affinity purification was repeated, allowing the ascites to circulate through the column for 1 hour before being washed, but did not change the elution profile. SDS-PAGE showed protein in the fractions eluting straight through the column only. A gel was run with both Blg and ascites fluid and then transferred onto nitrocellulose. Probing the nitrocellulose, primarily with ascites and secondly with anti-mouse conjugated HRP only showed the presence of the heavy and light chains from ascites. The antibody was therefore present within the ascites fluid but didn't appear to recognise Blg. The denatured state of the protein after its passage through a polyacrylamide gel may have influenced the monoclonals activity. A dot blot was used since the protein could be stuck directly to a sheet of nitrocellulose without its denaturation, and probed in an identical fashion to a western blot. After a buffer change from pH 9.6 to pH 7.0 the Blg was shown to stick to the nitrocellulose by means of an amido black stain. It is likely that at the higher pH, of 9.6, the protein is sufficiently negatively charged to be repelled by the nitrocellulose membrane. Incubation of the Blg immobilised on nitrocellulose with ascites fluid still showed no recognition of the protein from the antibody.

An ELISA using 10% glutaraldehyde, poly-L-lysine and 1% protamine sulphate independently as coating, to help the Blg stick to the plates, again indicated no recognition of the antigen from the antibody.

Monoclonals are highly specific proteins and it was possible that the Blg preparation may have influenced the binding pattern. Different preparations from SIGMA and PENTEX (Miles Laboratories) were examined using an ELISA and western blot. Neither was recognised by the antibody. alpha-Lactalbumin was also

investigated, in case the antibody was raised against the wrong milk protein. This also proved negative.

At this stage it was recognised that the antibodies were possibly inactive. An immunoprecipitation technique, using 50% saturated ammonium sulphate to precipitate the immune complex, was tried in case the antibody had a problem recognising the protein in an immobilised form. This was also unsuccessful since the unbound labelled protein was precipitated with a saturated ammonium sulphate concentration as low as 20%.

A polyclonal rabbit antiserum was raised towards Blg because of the failure of all previous attempts. This was to show that the previously used techniques were indeed working and that the Blg was being successfully immobilised and could be recognised by an antiserum. The rabbit was bled four times with each bleed being tested for activity in an ELISA. The serum showed maximal activity when diluted 1:50 and targeted against an antigen concentration of between 1 and 100µg/ml coated on the wells. The rabbit antisera recognised Blg in a western blot and antibodies were affinity purified by loading 4ml serum onto the immobilised Blg column. Protein bound was removed using 0.1M glycine-HCl (Figure 4.7), pH 2.0, analysed by an ELISA and showed greater activity than the original serum. Initial washings showed no activity (Figure 4.8).

Purified polyclonal antisera were used in a capture ELISA where the antibodies were coated on the plate prior to antigen addition. The antigen was then probed with the monoclonal antibody. The assay was shown to work by ^{125}I labelled rabbit antibody which confirmed the presence of the second antibody. Goat anti-mouse Ig HRP conjugate showed a negative result for the presence of the monoclonal.

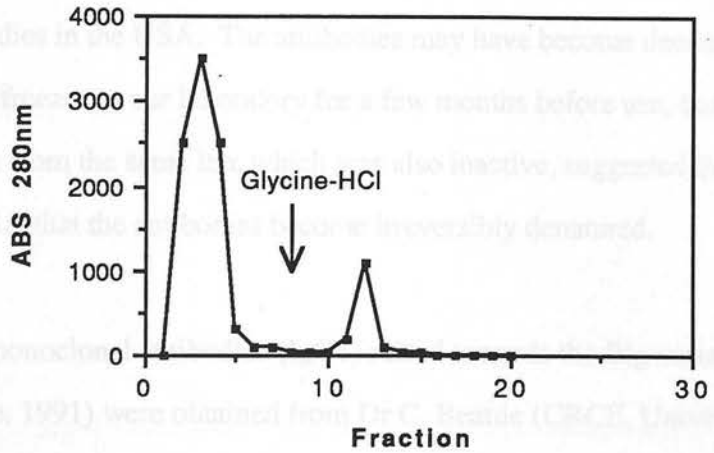


Figure 4.7 Rabbit antiserum elution profile from immobilised Blg column.

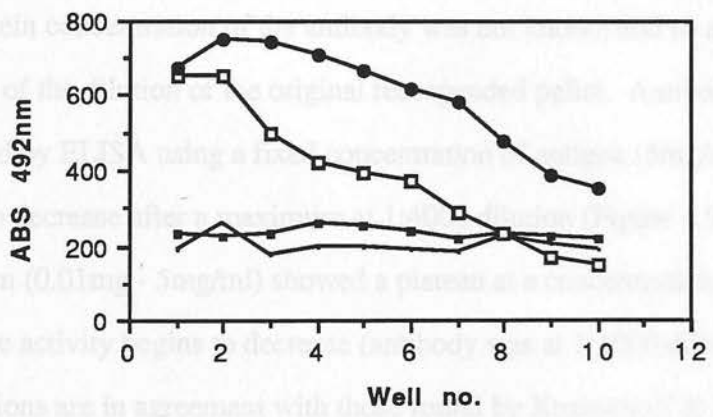


Figure 4.8 ELISA of affinity purified rabbit antiserum, before (□) and after (●) affinity purification. (■) pre-elution, (·) negative. Each well is diluted in half from the previous.

The final conclusion must be that both monoclonal antibodies are inactive. Procedures, indicated here, show little difference from those used to examine the antibodies in the USA. The antibodies may have become denatured by sitting in a -30°C freezer in our laboratory for a few months before use, but a second batch of ascites from the same lab, which was also inactive, suggested that it is in transit from the USA that the antibodies become irreversibly denatured.

New monoclonal antibodies (IgG1) raised towards the Blg variant A (Kuzmanoff & Beattie, 1991) were obtained from Dr C. Beattie (CRCE, University of Illinois, Chicago, USA). The antibodies were sent freeze dried and as an ammonium sulphate (AS) precipitate. The freeze dried preparation and the AS pellet (after spinning down) were resuspended in PBS. An ELISA showed no activity in the freeze dried material or the AS supernatant. Antigen recognition was, however observed in the resuspended AS pellet. This indicates that the state of the protein when it is sent is important in maintaining its activity.

The protein concentration of the antibody was not known and so activity is noted as a function of the dilution of the original resuspended pellet. Antibody avidity was examined by ELISA using a fixed concentration of antigen (5mg/ml). Activity was shown to decrease after a maximum at 1:4000 dilution (Figure 4.9). Serial dilutions of antigen (0.01mg - 5mg/ml) showed a plateau at a concentration of 1mg/ml, after which the activity begins to decrease (antibody was at 1:1000 dilution. These observations are in agreement with those found by Kuzmanoff & Beattie (1991).

The AS precipitated antibody therefore holds much of its activity unlike the other preparations that were tried. These antibodies, therefore, seem good candidates for preparing Fab fragments for the purpose of crystallisation of an immune complex. A further batch of monoclonal antibodies has been obtained and once these have been

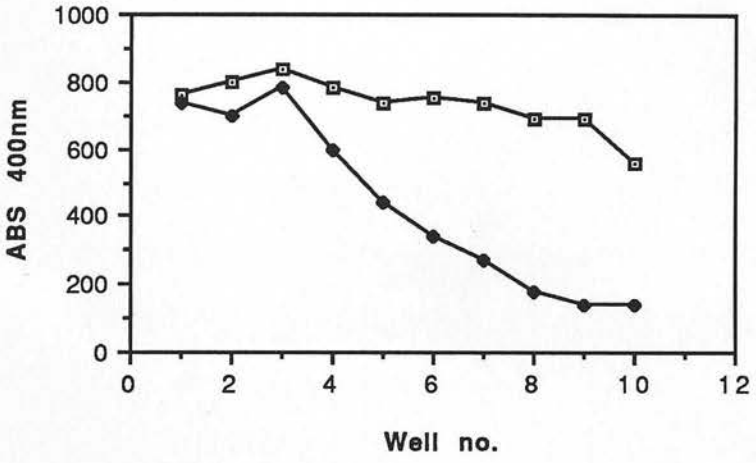


Figure 4.9 ELISA of mouse monoclonal antibody towards Blg. (⊡) antigen dilution 5-0.007μg/ml, (●) antibody dilution from 1:1000, halved along plate.

examined for their tight binding to Blg then work can begin into the preparation of a complex, and then onto the crystallisation of that complex.

CHAPTER 5

Crystallisation of ß-lactoglobulin.

Chapter 5

5.1 Introduction.

The advent of DNA technology has led to the availability of numerous proteins in quantities which allow more detailed studies to be carried out on them than would otherwise have been possible. However, despite the growing number of protein sequences that are now available, the number of corresponding structures that have been fully analysed is small. Often the rate limiting step in determining a protein structure is crystallisation of the macromolecule. Once crystals have been grown, finding a suitable form that will diffract to a reasonable resolution, can be difficult. Often several different crystal forms are required before a suitable native data set can be collected and analysed. Switching species can also be tried. As a result of these difficulties crystallisation methodologies are being investigated in a wide variety of ways, including zero gravity conditions, in the search for perfect crystals (Johnson, 1989). Crystals have been shown to grow, in some cases, to a higher quality than those grown under identical conditions on the earth.

A crystal is an ordered array of molecules, characterised by a unique set of parameters. The molecule, in this case a protein, gives up its many degrees of freedom to form stable chemical bonds with its neighbours in an ordered three dimensional object. Many crystals are regular in shape and exhibit a great deal of symmetry, though this need not always be the case. Crystallisation will only take place when repulsive interactions are minimised and attractive forces are maximised. This is achieved using highly concentrated protein solutions with the aid of an agent which preferentially hydrates the protein. Normally this agent is salt, which competes with the macromolecules for H_2O and encourages them to interact with one

another. Other agents, such as polymers or organic solvents, are thought to have a similar mode of action and can also be used. With respect to molecular weight, the number of interactions is few, causing the crystal to be delicate. Also, the positions of the molecules within different cells can vary slightly causing mosaic spread. However, the relatively few interactions in protein crystals tends to keep mosaic spread down.

The crystal, once grown, is surrounded by mother liquor. This liquid must be of sufficient ionic strength to allow the crystal to survive. In practice crystals are often transferred to a solution of higher ionic strength for storage. Various ligands, or heavy metal ions, may be added to the mother liquor and allowed to soak into the crystal via the solvent channels which exist.

Blg was first crystallised from bovine milk whey by Palmer (1934) as a result of the purification procedure. These crystals were later examined and revealed the dimeric nature of the protein (Crowfoot & Riley, 1939).

Some proteins crystallise in polymorphic forms, some of which can often be seen within the same batch. This is true of Blg. Hambling *et al.* (1991) indicate that at least 12 different crystal forms of Blg exist with lattices Y and Z growing under identical conditions (Green *et al.* 1979). The Y form is the more stable. Crystallisation under different conditions can aid in the investigation of some properties of the protein. Comparison of the lattice X (at pH 6.5) and lattice Y (at pH 7.8) structures of bovine Blg indicates an exposure of three carboxyl groups over the so-called Tanford transition (Hambling, 1990).

The crystallisation conditions of Blg are examined here with a view to using them in future structural investigation work. Where data have been collected, further

crystallisation attempts continue in an attempt to improve the reproducibility. Data collection procedures are outlined.

5.2 Crystallisation Methods.

In this project the main methods employed to crystallise Blg were the hanging drop, the batch method and that of microdialysis. These are described;

5.2.1 The Hanging Drop Method (Davies & Rich, 1971).

This technique employs vapour diffusion as a means of achieving the solubility minimum and obtaining crystals. Trials are conducted using a 24 well tissue culture tray (Flow laboratories) which provides discreet reservoirs into which buffered precipitant can be added. Commonly 5µl of protein solution is placed on a siliconised coverslip, diluted with 5µl of the reservoir buffer and then suspended over the reservoir and sealed using silicone grease. This method is useful as a means of establishing optimum conditions with a minimum of material.

5.2.2 The Batch Method (McPherson, 1982).

This technique is normally used once conditions have been narrowed to a reasonably small range. Normally the protein is added to a glass vial and the precipitant solution added to a concentration just below that at which the molecule precipitates. The vials are then sealed and left to stand for inspection at a later date.

5.2.3 Microdialysis (Zeppezauer *et al.*, 1967).

In this technique small volumes of protein (commonly 20-100µl) are placed in a small button which is sealed off using dialysis membrane. The button is then placed in a beaker containing buffered precipitant and left to stand. Equilibration occurs across the membrane and crystallisation can occur. Slower equilibration can be achieved by double dialysis, where the protein is separated from the most concentrated precipitant solution by a second membrane. The advantage of this technique is the concentration of precipitant can be easily changed.

5.3.1 Crystallisation of a Blg-Sodium Dodecyl Sulphate Complex.

The use of detergents in the crystallisation of macromolecular structures is quite rare. Crystals have been grown in the presence of such compounds in the cases of purple membrane protein (Michel & Oesterhalt, 1980) and the *E.coli* plasma membrane protein (Garavito & Rosenbusch, 1980). McMeekin *et al.* (1949) recognised that at pH's close to a protein's isoelectric point neither anionic nor cationic detergents will form a precipitate and succeeded in crystallising a complex of Blg with sodium dodecyl sulphate (SDS). Considering the two previously mentioned proteins are integral membrane proteins it seems quite unusual that Blg should crystallise when complexed with SDS.

Crystals of the complex were obtained here by the addition of 100µl 0.01M SDS to 250µl Blg solution at pH's between 4.6 and 5.0 (50mg/ml in 10mM citrate phosphate buffer). This modification of McMeekin's procedures only succeeded in obtaining small rectangular plate crystals covered in precipitate. For this reason they were not immediately apparent until they were shown to shine under polarised light. Crystals

took two days to grow and any attempt to move them resulted in them shattering. As a result of the fragility and very thin nature of the crystals they were unsuitable for X-ray work.

The use of different protein preparations influenced crystallisation. Crystals were only obtained on one occasion and from a preparation from PENTEX (Miles Laboratories). Neither protein from SIGMA or BDH yielded crystals. Further trials, where conditions such as temperature, pH, and protein concentration were varied, had no effect. Button dialysis of 50mg/ml Blg against 0.01-0.1M SDS did not result in crystal formation.

If future trials yield suitable crystals then the interaction of SDS with Blg will be elucidated. Both McMeekin *et al.* (1949) and Seibles (1969) indicate that 2 molecules of SDS bind per dimer of Blg. Seibles proposes that His 146 is essential for the interaction. On examination of the lattice Y model it appears that a single SDS molecule may lie in a cleft close to the α-helix. Met 145, Met 24 and the free cysteine 121 may play a role in holding the ligand in position (Figure 5.1). These residues are also mentioned by Seibles, though he emphasises their importance to a lesser extent. The binding site is close to that seen for retinol in the lattice Z model (Monaco *et al.*, 1987, Figure 5.2), though the SDS may lie deeper in the pocket. On unfolding the protein will create many other potential ligand binding sites and still hold onto the SDS in the pocket it already occupies (Seibles, 1969). Under denaturing conditions, e.g. in gel electrophoresis, where SDS is used,presumably the mode of action is somewhat similar. SDS molecules eventually engulf the protein conferring an overall charge on the protein.

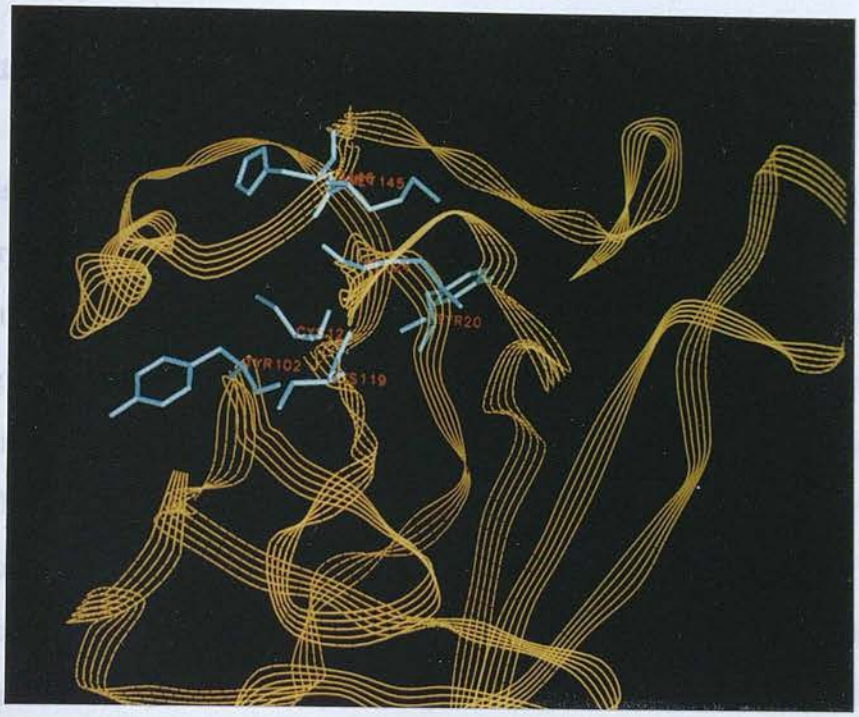


Figure 5.1 Picture showing cleft for SDS binding to Blg. The residues proposed to be involved are outlined.

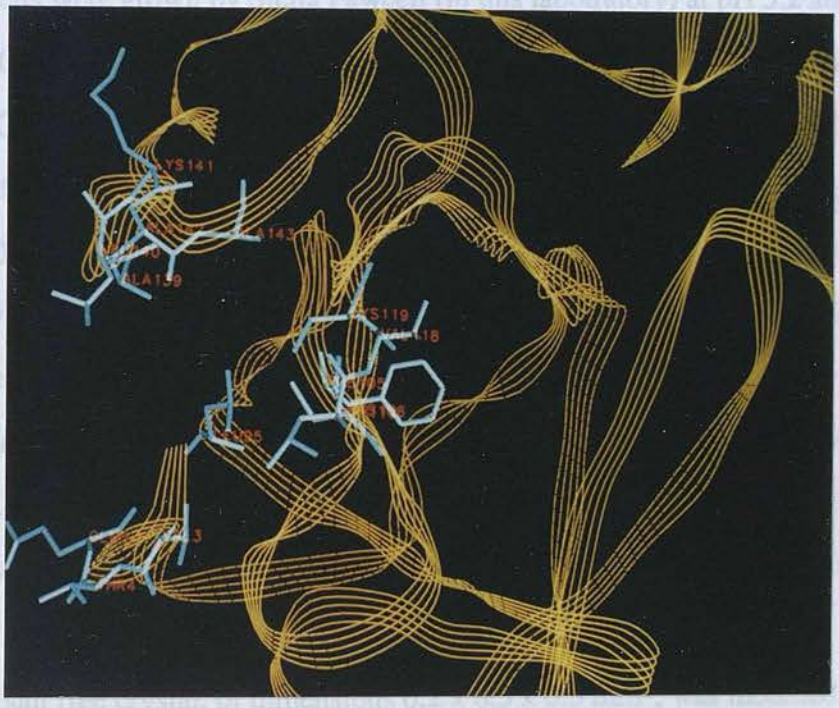


Figure 5.2 Picture showing the external binding site for retinol. Residues in close contact as described by Monaco *et al.* (1987) are outlined.

5.3.2 Salt Free Crystals.

Crystals grown from a salt free medium can have advantages over their salt grown counterparts. Polyethylene glycol (PEG) is often used on such occasions. It is thought that the polymer acts as a hybrid between the effects of salt and alcohols, although the true mechanism is not entirely understood. Often crystals of specific proteins grown from PEG can exhibit the same symmetry and dimensions as those grown by other means (McPherson, 1982). The advantage of such crystals is that they can be grown from lower ionic concentrations, which allows easier incorporation of heavy atoms and ligands into the mother liquor for use in soaking experiments. Also, a low electron density medium is provided by PEG which can prove helpful during map interpretation at a later stage (McPherson, 1976)

Blg crystals were grown by Stella Fawcett (in this labotratory) at pH 5.2 and 5.6 from a 50mg/ml solution of protein, obtained from the Hannah Institute (Ayr), using 12% PEG 4000 as precipitant in 10mM acetate buffer. All trials contained 0.1% azide as fungal growth occurred rapidly in its absence. In this project trials using industrially prepared protein, SIGMA and PENTEX (Miles Laboratories), failed to grow crystals by both the hanging drop and the batch methods. It is possible that the crystallisation of the protein under these conditions is highly sensitive to the protein preparation, or that the conditions of initial crystal growth have not been suitably matched. It seems unlikely that it is the latter, since various conditions, around which crystals were originally grown, were exhaustively tried and yielded only precipitate.

A single salt free crystal, of dimensions 0.2 x 0.3 x 0.3 mm^3, was mounted in a Lindemann tube before being examined by precession photography. A crystal was launched into the mother liquor (20% PEG) and allowed to drop to the lower miniscus within the tube, before being dried out. The mother liquor was centrifuged

before being allowed to come into contact with the crystal. If it was not spun then the PEG came out of solution around the crystal giving an unclean crystal mounting. Precession photography indicated the space group to be $P2_1$ with cell dimensions of a = 72.2 Å, b = 67.9 Å, c = 36.2 Å, ß = 92.0°. Data were collected on these crystals using a STOE AED4 diffractometer up to 5.0 Å using Cu-Kα radiation (λ=1.54 Å). The positions of 12 reflections, collected between 20° and 40° in 2 θ, were used to determine the orientation matrix and check for slippage during data collection. Absorption corrections, derived from Psi scans were applied prior to data processing. Due to problems during data collection the data were not of a high quality. Chapter 6 investigates the possibility of using the technique of molecular replacement to give initial phases for this structure. Higher resolution data needs to be collected and added to the 2096 reflections already collected.

The crystal form shows remarkable resemblance to that of the lattice W, found by Aschaffenburg *et al.* (1965), except that the length of cell edge a is half of that found here. The lattice W form was grown by dialysis against water. The crystal form is a derivative and contained either PCMS or iodoacetamide. Aschaffenburg also found that crystals were unstable in the X-ray beam, a factor which was not evident here, though diffraction did not extend beyond 5.0 Å.

It is clear that this crystallisation technique is sensitive to the preparation of the protein. Collection of data to a higher resolution has been hindered by the lack of a suitable batch of protein, and therefore crystals.

5.3.3 Crystallisation of Blg at Low pH.

With the exception of precipitant, pH is seen as the most important variable in

crystallisation (McPherson, 1982). Proteins, by virtue of their polyionic nature, can vary dramatically in solubility over a wide range of pH and Blg is no exception. It does, however, exhibit a remarkable stability at low pH (Aschaffenburg & Drewry, 1957). Armstrong *et al.* (1967) modified this purification procedure, using ammonium sulphate fractionation, as a means of separating the protein without lowering the pH to levels which would dissociate the protein.

It is thus worth examining the bovine protein at pH's below 3.5 in its crystalline form to ascertain whether any changes have occurred as a result of monomerisation (Townend *et al.*, 1960). Studies by CD (Townend *et al.*, 1967), ORD (Timasheff *et al.*, 1966) and Trp fluorescence (Townend *et al.*, 1967) indicate no major conformational changes have occurred. Also, as a result of interest into the use of the protein as a drug carrier (McAlpine & Sawyer, 1990), it is important to know the low pH structure which may differ from others in a significant way. The soaking of prospective ligands will be facilitated if a crystal form at low pH is available

Crystals were originally grown by the hanging drop method between pH 3.0 and 3.3 using 15mM citrate phosphate buffer. The protein was a mixture of bovine Blg variants A and B (SIGMA) at 50mg/ml in buffer at pH 3.5, and used ammonium sulphates precipitant at 2.0 to 2.6M. Crystals grew to 0.2 x 0.2 x 0.1 mm^3 (Figure 5.3) though were generally covered by precipitate (Figure 5.4). Lowering the protein concentration produced fewer crystals. Dissolving the protein in H_2O, rather than buffer, resulted in showers of very small crystals which were of little use. An addition of 100µl of protein at pH 3.5 to 200µl of buffer at pH 3.0 to 3.3, using the batch method, allowed crystals to form after one month in one batch only. At pH 3.1, with an ammonium sulphate concentration of 2.2M, crystals were obtained as large hexagonal plates, dimensions 0.3 x 0.3 x 0.15 mm^3, or as small hexagonal rods.

Figure 5.3 Crystals grown at pH 3.1 from 2.0M ammonium sulphate by the hanging drop method. Approximate dimensions of 0.2 x 0.2 x 0.1 mm^3.

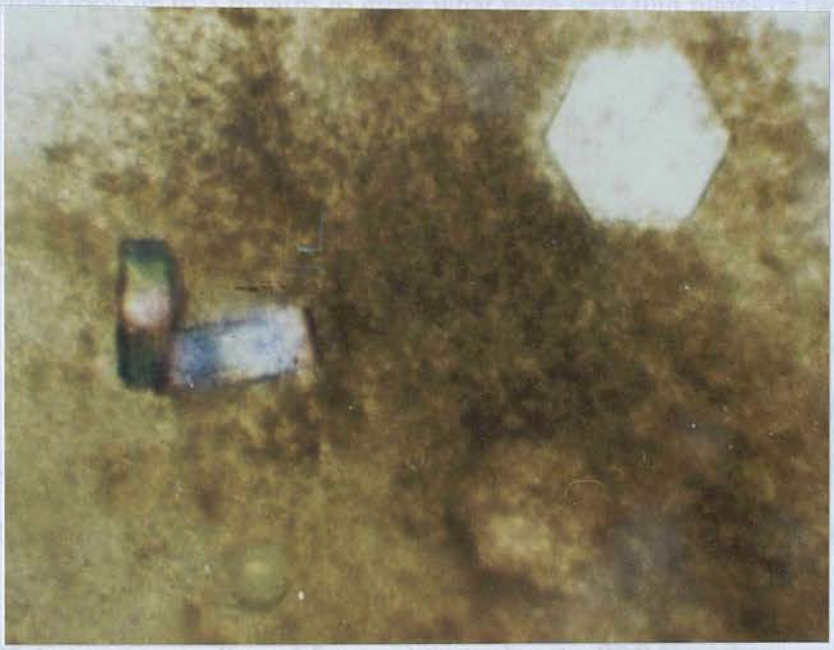


Figure 5.4 Crystals grown by the hanging drop method as in Figure 5.3. Picture shows protein precipitate around the crystals.

Any change to the crystallisation conditions was detrimental to crystal growth. Dissolving the protein in buffer at a pH other than 3.5, did not allow crystal formation, as did any change in ionic strength of the buffer. After many failed trials another factor was found which seemed to influence crystal growth. Buffers, which was lowered to the correct pH using 1M HCl, rather than any other acid, allowed the growth of crystals by the hanging drop method.

The hanging drop seems to be most reliable at reproducing crystals with the conditions being optimised at an ammonium sulphate concentration of between 1.90 and 2.15M at pH 3.0 to 3.3. However, these crystals show a considerable disorder in diffraction experiments. The effect of temperature on crystal growth was also examined and indicated the best crystals, as far as morphology was concerned, at 20 °C, with crystals also growing to a lesser extent at 10 °C. Higher and lower temperatures did not allow crystal formation. The batch method was unreliable and on occasions took up to 3 months to grow what were very thin plates or twinned crystals, neither of which were of any use in diffraction experiments (Figure 5.6)

Crystals from the first batch grown, were examined by precession photography. Photographs of the h0l and hk0 regions indicated cell dimensions of a = b = 68.49 Å, c = 143.17 Å, and $\gamma = 120°$. Crystals were from the hexagonal space group $P6_3$. These dimensions resemble, though are not identical, to those found by Steinrauf (1959) but are grown under different conditions. Bolognesi *et al.* (1979) indicate another hexagonal form with similar cell parameters from pooled milk of the buffalo.

A full 60° data set was collected from one crystal at the SRS Daresbury, station 7.2, to 3.3 Å. The crystal showed an extremely high degree of mosaic spread with stills bearing strong resemblance to precession photographs. Closer examination of the

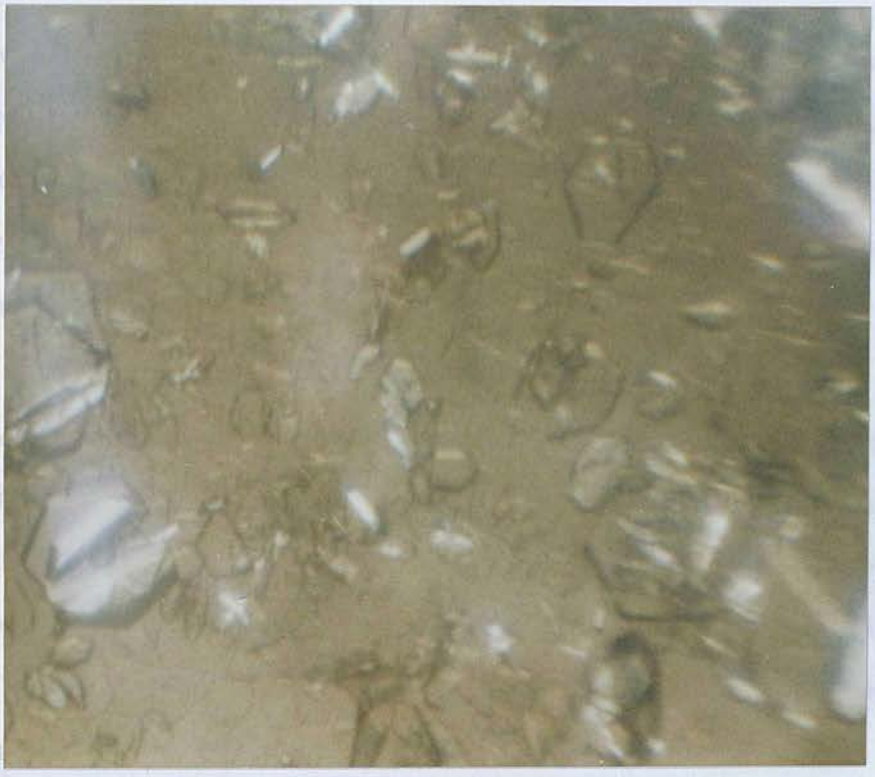


Figure 5.6 Crystals of Blg grown by the batch method. The picture shows the crystals to be plate-like or badly twinned.

crystals shows them to be very plate like, i.e. each crystal is made up of numerous hexagonal plates stuck on top of one another. 30° of high resolution data was collected to 2.9 Å from 4 crystals. Crystals tended to die rapidly due to the longer exposure when attempting to obtain high resolution data.

These data have still to be processed. Meanwhile the search for more reliable crystals with less mosaic spread is being investigated.

5.4 Conclusions.

The major factors that come out of all of the trials examined here is the importance of a good protein preparation and the necessity to reproduce conditions exactly. What is interesting is that the same protein preparation did not always give crystals at different conditions, e.g. the complex with SDS and the low pH crystals were from different protein preparations.

Crystallisation over varied conditions will ultimately lead to a more comprehensive insight into the proteins structure and may even lead to a firm function being assigned to the protein. It will show conformational changes over a range of pH values and give an insight into the binding properties of the protein. Comparisons of the lattice X (Yewdall, 1988) and lattice Y (Hambling, 1990) structures have already given an idea as to what happens structurally over the Tanford transition.

The use of ligands may ultimately produce crystals of a better quality, because of stabilisation set up by interactions with the ligand. Crystals of the bovine Blg lattice Z complexed with para-nitro phenol show stronger diffraction than native crystals but are non-isomorphous, having a contraction of the c-axis (Morais-Cabral, personal communication).

If the protein is to be used as a drug carrier then its structure at low pH will be of importance. The data already collected have still to be processed but should give some insight into some important residues in this pH region. Crystallisation conditions need to be further examined, however, if the effect of mosaic spread is to be reduced.

CHAPTER 6

Refinement & Molecular Replacement.

Chapter 6.

6.1 Introduction.

The solution of two small molecule structures was described in chapter 2. This served as an introduction to the crystallographic process and allowed the transition to protein crystallography. The following chapter describes the current state of refinement for the Blg, lattice Y, model structure. It follows the process to determine a more accurate description of the Blg structure, the problems encountered, and the steps taken to overcome them.

The latter part of this chapter describes the attempts at locating the Blg molecule in another crystal form. The technique of molecular replacement is employed in initial attempts at finding the orientation of the unknown Blg molecule using the Blg lattice Y structure as a search molecule.

6.2 Refinement.

Phasing errors lead to an initial electron density map and hence a molecular model which is inaccurate. Refinement produces for the crystallographer a more accurate description of the protein. The various parameters are adjusted to give the best agreement between the observed structure factor amplitudes and those calculated from the model by the following equation:

$$Fc_{(hkl)} = {}^{N}\Sigma_{j=1}\, P_j\, f_j \exp\left[-B_j\,(\sin^2\theta/\lambda^2)\right] \exp 2\pi i\,(hx_j + ky_j + lz_j) \qquad (1)$$

N = number of atoms

P = occupancy (normally = 1)

f = atomic scattering factor

B = temperature factor

x,y,z = atom coordinates

Refinement of atomic models is rapidly becoming a standard aspect of crystallographic studies on macromolecules and generally takes either of two approaches.

Real space refinement makes use of electron density maps and the best fit for the atoms within this map are found. Difference electron density maps are calculated using the difference between the observed and calculated structure factor amplitudes, phase information is taken from the model. Atoms missing from the structure show up as peaks and wrongly placed atoms show up as holes, i.e peaks with negative density. This can help the crystallographer to best choose atom positions for his structure.

Reciprocal space refinement is more common. The inaccuracy of initial models and the lack of observable data is overcome by the prior knowledge about the stereochemistry of macromolecules. Least squares techniques are used to minimise the quantity in equation 2. The method of least squares is powerful technique allowing the best fit of a mathematical model to a set of data. Restrained least squares is reviewed by Hendrickson & Konnert (1981) and involves the minimising of equation 2 such that the number of observations is increased so the problem does not become underdetermined (i.e. the number of observations should exceed the number of variables).

In this case the package TNT (Tronrud *et al.*, 1987) was used to minimise the following equation:

$$M = \Sigma_j W_j [Q_{o(j)} - Q_{c(j,p)}]^2 \qquad (2)$$

$Q_{o(j)}$ = the value of observation j

$Q_{c(j,p)}$ = the value calculated from model coordinates and thermal parameters (p)

W_j = a weighting factor

The equation is summed over all observations and can be split into stereochemical, (b), and crystallographic observations, (s):

$$M = \Sigma_s W_s [Q_{o(s)} - Q_{c(s,p)}]^2 + \Sigma_b W_b [Q_{o(b)} - Q_{c(b,p)}]^2 \qquad (3)$$

The package utilises restrained least squares and is fast and flexible, as a result of the way in which it has been organised. The program splits each set of functions into modules which it can treat separately. Each module calculates the quantity for its term, after which the control program collects a vector gradient from each module and determines the direction and the magnitude of the shift to each parameter. A flow diagram of the modules within the package is outlined in Figure 6.1.

The crystallographic module encompasses five separate programs which allows a specific fast Fourier transform (FFT) to calculate the structure factors. This FFT is space group specific.

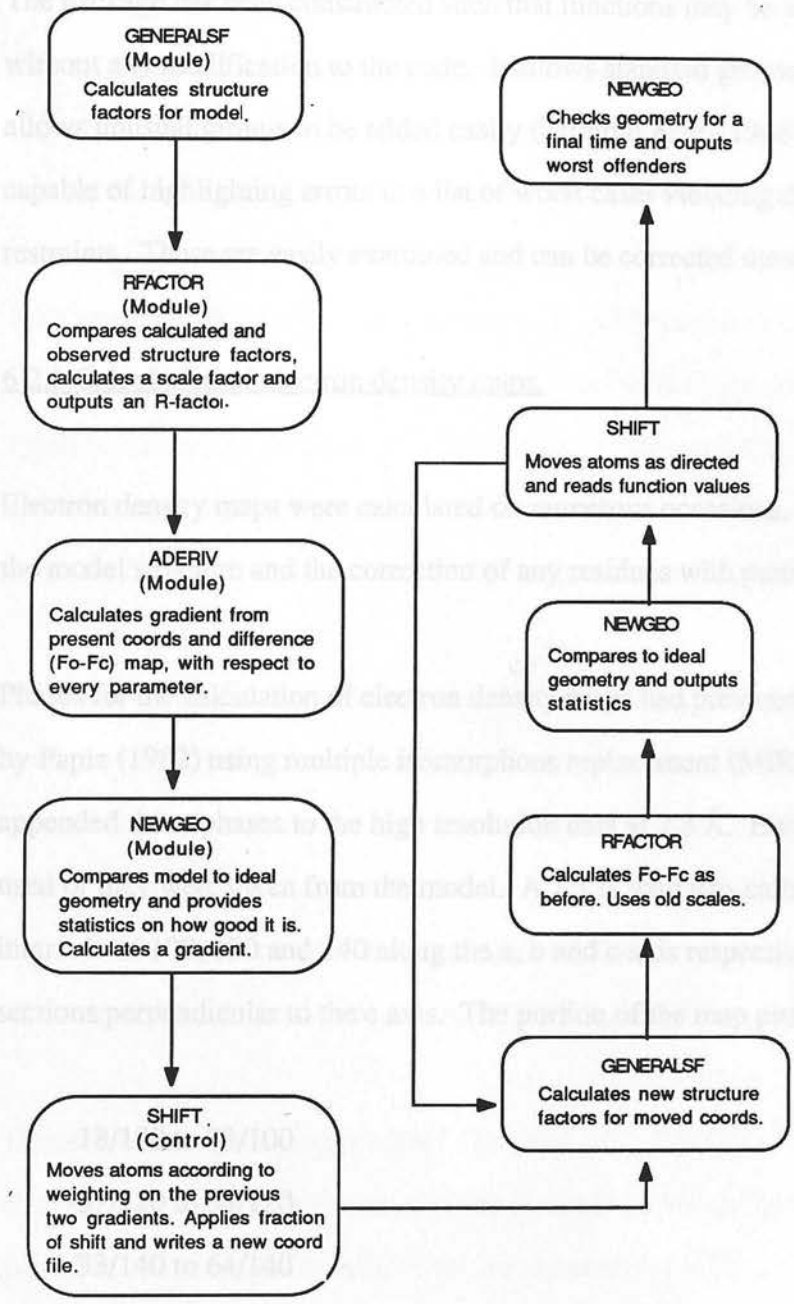


Figure 6.1 TNT flow diagram. Shows the steps taken to complete one refinement cycle.

Data input and output are free format which allows ease of transfer between modules. The package has been constructed such that functions may be added, or replaced, without any modification to the code. It allows standard geometry to be defined and allows unusual groups to be added easily (Tronrud *et al.*, 1986). The program is also capable of highlighting errors in a list of worst cases violating the refinement restraints. These are easily examined and can be corrected manually.

6.2.1 Calculation of electron density maps.

Electron density maps were calculated on numerous occasions, to allow inspection of the model structure and the correction of any residues with particularly bad geometry.

Phases for the calculation of electron density maps had previously been determined by Papiz (1982) using multiple isomorphous replacement (MIR). Hambling (1990) appended these phases to the high resolution data at 2.8 Å. Either these phases were used or they were taken from the model. A 2.8 Å map was calculated with sampling intervals of 100, 120 and 140 along the a, b and c axis respectively and was in sections perpendicular to the c axis. The portion of the map produced was

-18/100 to 73/100
-37/120 to 58/120
-33/140 to 64/140

This was displayed using the program FRODO (Jones, 1978) on either an Evans and Sutherland PS300 or ESV20.

6.2.2 Initial refinement of Blg.

The original model of Blg was of a coordinate set which had been rebuilt into a 2.8 Å (2Fo-Fc) electron density map by S. Hambling (1990). A brief description is given here of the work that had been carried out on the structure up to this stage.

Low resolution data collection was originally performed by Green *et al.* (1979) to 6 Å. This was extended to 2.8 Å resolution data by M.Z. Papiz (1982). Two medium resolution data sets, one film the other diffractometer, and the 6 Å data were combined and the original model chain traced in a 2.8 Å electron density map which was calculated using phases derived by the method of isomorphous replacement.

High resolution data were collected to 1.8 Å resolution (Hambling, 1990) and merged with the sets already mentioned. These data, phased to 2.8 Å, were used to calculate a map which was deemed to be better than that generated by Papiz (1982), although still showed little continuous main chain density for residues 55 to 69 and 150 to 162. Each of these data sets were collected under different systems and may be the cause of some of the problems encountered in the refinement.

Here, the refinement program TNT (Tronrud *et al.* 1987) was used, employing restrained least square refinement to reduce the R-factor and give acceptable geometry. The stages of refinement are summarised in Figure 6.2. The starting atomic coordinates had an initial R-factor of 47.8%. A series of refinement cycles, to examine the effects of varying different parameters helped to reduce the R-factor to 44.7% with no adverse effects on the geometry of the model. Refinement cycles which reduced the R-factor to 34.0%, with a high weight on the X-ray terms, were discarded because of adverse effects on the geometry of the model.

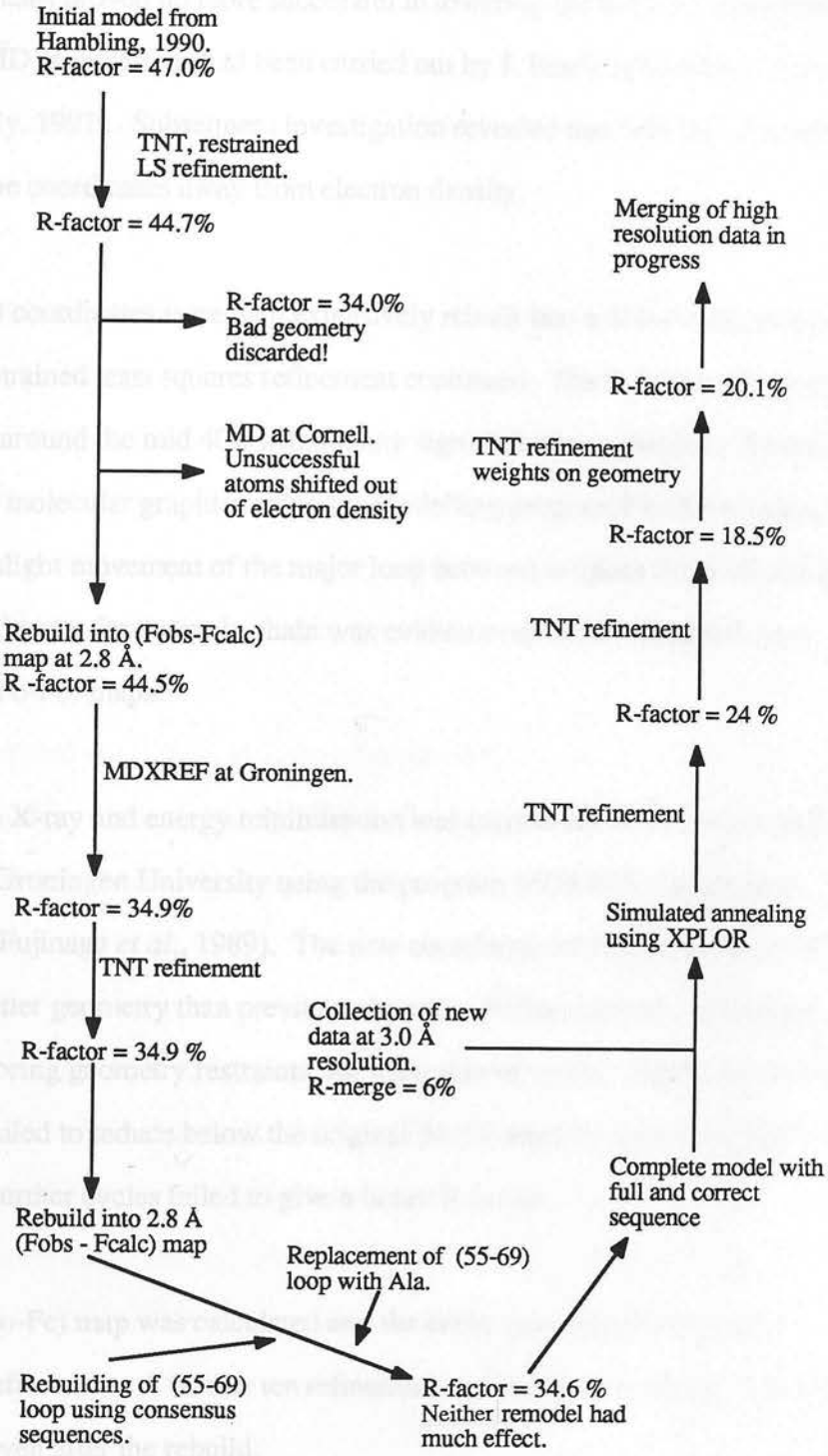


Figure 6.2 An overview of the refinement of Blg lattice Y.

The coordinates proved no more successful in lowering the R-factor after molecular dynamic (MD) simulations had been carried out by J. Brady at Cornell University (Gu & Brady, 1991). Subsequent investigation revealed that MD had only succeeded in pulling the coordinates away from electron density.

The original coordinates were then extensively rebuilt into a 2Fo-Fc electron density map and restrained least squares refinement continued. The R-factor still proceeded to fluctuate around the mid 40's without any sign of further reduction. Studying the structure on molecular graphics using the modelling program FRODO (Jones, 1978) indicated a slight movement of the major loop between residues 55 to 69, though still no electron density for the main chain was evident even when examined with difference (Fo-Fc) maps.

At this stage X-ray and energy minimisation was carried out by L.Sawyer and M. Fujinaga at Groningen University using the program MDXREF, based upon GROMOS (Fujinaga *et al.*, 1989). The new coordinate set had an R-factor of 34.9% and had a better geometry than previous attempts. Refinement was continued using TNT, monitoring geometry restraints, for a number of cycles. The R-factor fell to 32.9% but failed to reduce below the original 34.9% once the geometry had been corrected. Further cycles failed to give a better R-factor.

An Fobs (2Fo-Fc) map was calculated and the entire structure rebuilt before continuing refinement. A further ten refinement cycles failed to reduce the R-factor any further even after the rebuild.

The main worry at this stage was the contribution to the model structure factors from a loop which had no apparent density (residues 55 to 69). To reduce the effect of the side chain atoms in this loop the residues between 50 and 70 were converted to

alanines using the REPLACE option within the module SAM of FRODO. Eight refinement cycles reduced the R-factor to 29.5%, but returned to 34.6% once the geometry terms had been refined. The loop residues were replaced since changing them seemed to have little effect.

6.2.3 A New Low Resolution Data Set.

As a result of the problems with the model, and the amount of time which was required to reduce the R-factor, it was decided to collect a new medium resolution data set. The set at present was merged from 4 different sources. It is likely that part, at least, of the problem lay in this set of low to medium resolution reflections.

A lattice Y crystal was mounted in the usual manner and two data sets were kindly collected from the single crystal by Elspeth Gorman in Oxford. The data , collected on a Xentronics area detector, was calculated to be 90.5% complete to 3.0 Å and had an R-merge of 6%.

The data set contained 3032 reflections, 3026 of which were common with the original data that was being used. The program R-Factor (CCP4) calculated a scale factor between the relevant reflections from a Wilson Plot and showed the R-factor to be 13.60% using the structure factor amplitudes from the new data set as Fobs (scales were applied to Fcalc). This gave a good indication that there were some considerable differences between the two data sets.

The most recent model, along with the new 3.0 Å data was then subject to molecular dynamics refinement in XPLOR (Brunger *et al.*, 1987) with the help of P. Adams.

6.2.4 Molecular Dynamics.

Traditional refinement techniques require that the model be remodelled into electron density, a step which is time consuming. Restrained least squares will not correct residues, or atoms, which are displaced by greater than 1 Å. The process becomes easily trapped in local minimum and human intervention is, therefore, required. Molecular dynamics utilises the potential energy that each atom has, calculated from Newtons equations of motion (Verlet, 1967), and allows larger areas of conformational space to be searched.

Here the program XPLOR (Brunger *et al.* 1987) was used to search for a more accurate model of the Blg structure using the data set that was collected to 3.0 Å in Oxford.

A flow diagram of the program can be seen in Figure 6.3. From the model coordinates (with H's added), and a library of the stereochemistry of amino acids, a topology file is generated. All charges are removed from side chains at this stage and the process is carried out "in-vacuo". XPLOR calculates the initial weights that will be applied to the X-ray terms as a result of minimising the structure at 300K. These weights are applied and the X-ray term is recalculated if the atoms are seen to move by greater than 0.05 Å. The system is heated to 2000K and atom velocities recalculated every 25 of 2000 steps (1 step = 1 fsec). The molecule is now at equilibrium at 2000K. The system is then rapidly cooled and equilibrated at 300K with velocities being recalculated every 25 of 250 steps. Finally the system is again energy minimised to remove any peculiar geometry that has arisen and is ready to use in the subsequent stages of refinement.

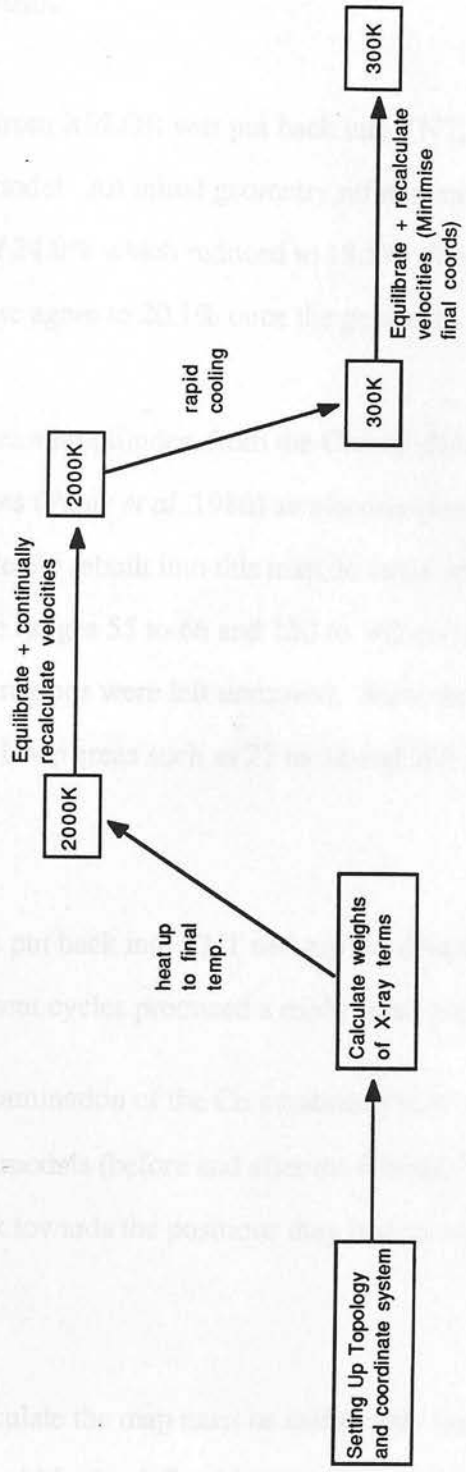


Figure 6.3 XPLOR flow diagram. Model weights are calculated and the atoms heated to 2000 Kelvin. After equilibration the molecule is cooled and the structure minimised.

6.2.5 Further Refinement.

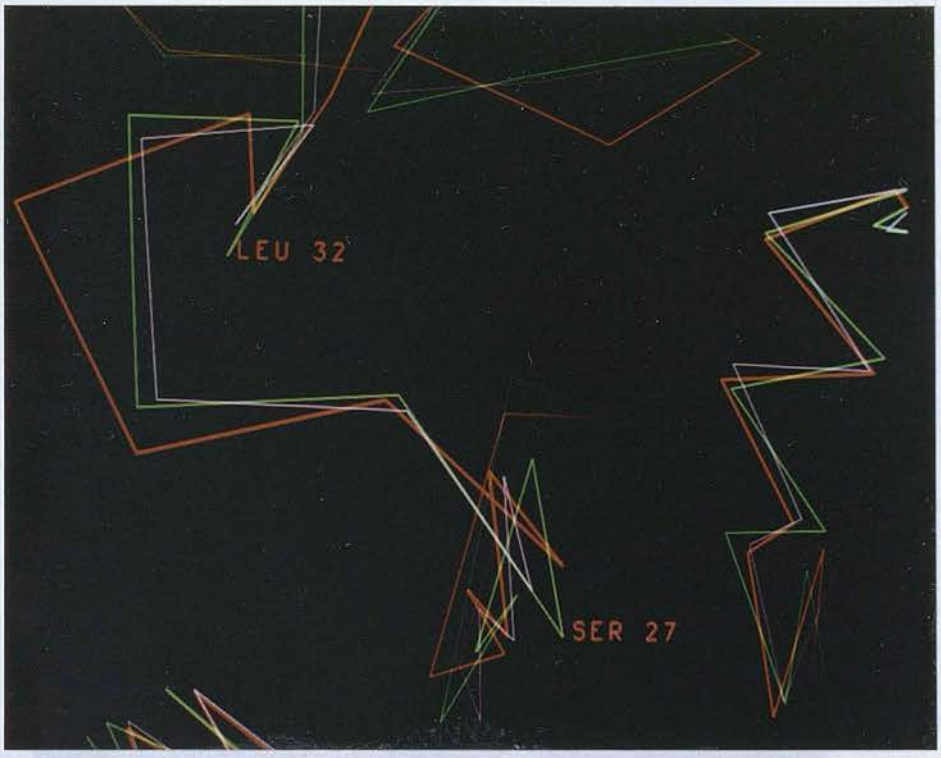
The resulting model from XPLOR was put back into TNT, after removing the hydrogens from the model. An initial geometry refinement using the new data to 3.0 Å gave an R-factor of 24.0% which reduced to 18.5% after 6 restrained least square cycles. This value rose again to 20.1% once the geometry had been optimised.

Using the structure factor amplitudes, from the Oxford data, and the initial MIR, solvent flattened phases (Papiz *et al.*,1986) an electron density map was generated. The model was completely rebuilt into this map, in areas where main chain density existed. In the residue ranges 55 to 66 and 150 to 162 no main chain density was apparent and so these regions were left unmoved. Some areas were quite extensively moved to fit the map. Loop areas such as 27 to 32 and 109 to 115 being the most apparent (Figure 6.4).

The rebuilt model was put back into TNT and proved disastrous. The R-factor was 44.9%. Three refinement cycles produced a model with poor geometry and an R-factor at 39.1%. Examination of the Cα positions produced at this stage with those from the previous two models (before and after the rebuild) indicated that the atoms were being pulled back towards the positions they had occupied before the rebuild (Figure 6.5).

The phases used to calculate the map must be sufficiently inaccurate to describe the trace of this molecule, within the defined loop regions anyway. It was decided to continue with the data set to 3.0 Å and the model prior to the rebuild. The high resolution data could be rescaled and and then merged to this new data set.

(a)



(b)

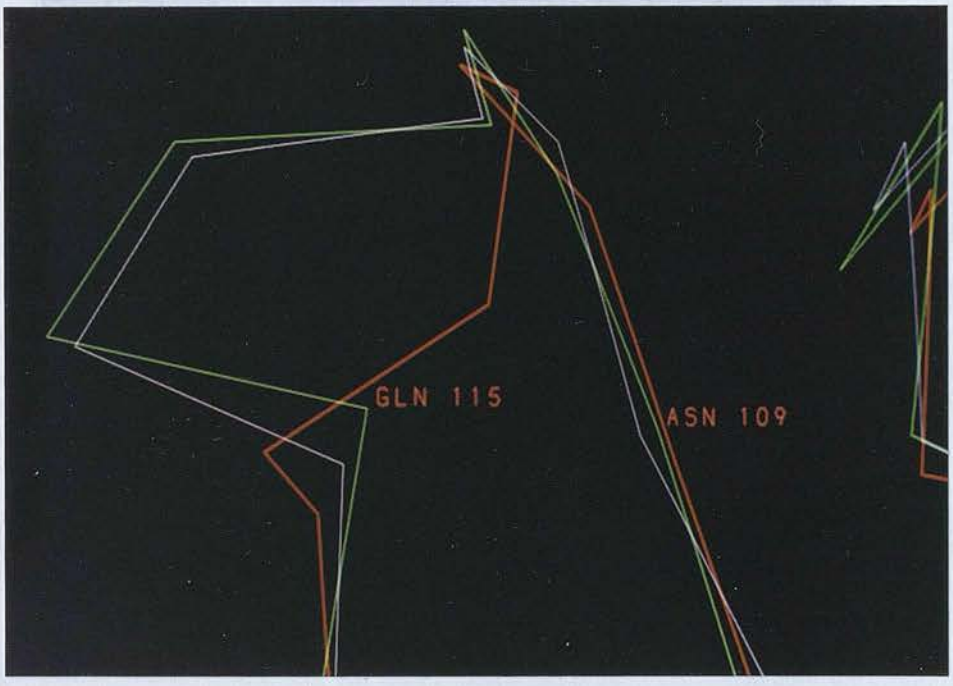


Figure 6.4 Regions of movement in the lattice Y structure before (white) and after the rebuild (red) into the MIR map (Green is after TNT refinement). (a) is loop 27-32; (b) is loop 109-115.

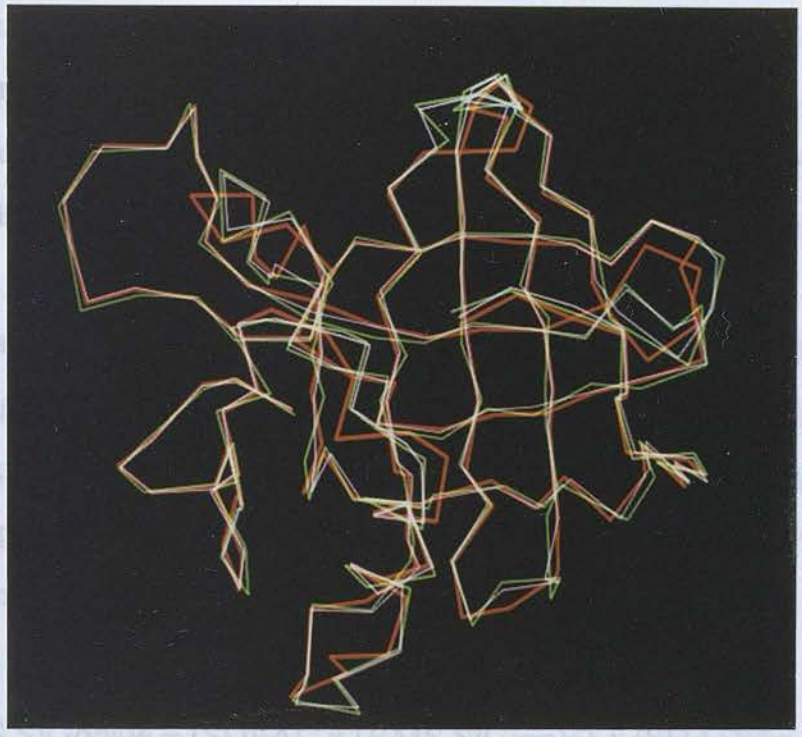


Figure 6.5 Picture showing movement of the loop regions of the Blg lattice Y model. Regions shown are: before the rebuild into the MIR map (white), after rebuild (red), after TNT refinement (green).

6.2.6 Scaling and Merging High Resolution Data.

Data collected at the SRS at Daresbury to 1.8 Å resolution were collected by S.Hambling (1990). The data had been previously scaled, but it was decided to go go back to the raw filmpack data and rescale in the hope of removing any errors that may have occurred originally.

The data were collected from 5 crystals and each crystal set contained batches of scaled film packs. For each crystal, individually, initial scale factors between batches were calculated by taking the ratio of the mean intensity between overlapping reflections in each batch. The scales were then refined by the method of Fox & Holmes (1966). Table 6.1 shows a summary of the scaling of each crystal. Partially recorded reflections were summed and symmetry equivalents were averaged.

The data from each of the 5 crystals were then merged before scaling them together. Scaling did not give adequate values for the mean and sigma for the full reflections and gave a merging R equal to 20.8%. Omitting the worst scaled crystal set (Crystal C) helped to reduce the merging R to 13.4%. Using SDADD and SDFAC in the equation

$$(\text{RMS SIGobs})^2 = (\text{SDFAC} \times (\text{RMS SIGinp})^2) + (\text{SDADD} \times I^2) \quad (4)$$

allowed reasonable values for the mean and sigma, of 0.15 and 0.96 respectively, to be obtained.

To check that the unique data did not diminish in a particular area the completeness of the data sets were examined both with and without the data from crystal C present.

Crystal	A	B	C	D	E
Reflections Input	28,235	15,162	21.788	14,253	13,438
Reflections used	22,997	11,983	18,891	10,952	11,206
No. full reflections	21,457	11,442	18,238	10,417	10,919
No. partial reflections	1,540	541	653	535	287
Mean Intensity	423	1,668	615	953	645
Mean SD	-0.01	0.02	0.24	-0.01	0.12
Independent HKL's	8,026	5,107	5,898	4,202	3,947
Reflections output	8,933	6,640	6,618	4,715	4,166
% > 3SD	57	64	72.9	71.5	75.4
R-merge	0.08	0.062	0.21	0.048	0.074

Table 6.1 Analysis of the scaled data from each of the bovine Blg crystals.

Table 6.2 shows that removal of the data does diminish the amount of unique data but it is over a range of d values and, therefore, should not have such a noticeable effect.

6.2.7 Discussion.

The scaling and merging of the new low resolution data to the high resolution set is underway and should give a complete and unique data set to 1.8 Å. These data can then be examined in TNT and model refinement continued. Phases will be calculated by the method of isomorphous replacement. A platinum derivative of a lattice Y crystal has already been collected on a Xentronics area detector at Glasgow University (Dept. of Chemistry), but has still to be examined against the native. Once phases have been calculated refinement can continue. Phases will probably be extended to the higher resolution using XPLOR.

At the present stage the structure will benefit from the introduction of the high resolution data and accurate phase information. The R-factor is at a value lower than it has been at previously. However this lower R-factor seems to be at the expense of less accurate model geometry. Ramachandran plots at three stages in the refinement process can be seen in figure 6.6. Figure 6.6 (a) shows the phi psi relationship for the starting model at the beginning of this study. Figure 6.6 (c) shows a movement of the values towards the more allowed regions of space from the final model discussed here. Figure 6.6 (b) shows a ramachandran plot after the MD using GROMOS and then the rebuild into the 2Fo-Fc electron density map and has the best distribution of values.

An omega plot of the final model (Figure 6.7) shows three residues to be in the cis-peptide conformation. Examination of the model indicates that each of these

Resolution (Å)	Percentage of unique data	
	+C	-C
5.00	61.7	57.1
3.08	88.3	83.2
2.56	93.6	89.1
2.23	95.3	91.1
2.01	96.3	91.9
1.84	94.5	90.1
1.71	17.7	17.6

Table 6.2 Percentage of unique data present, as a function of resolution, in the combined data sets with crystal C both absent and present.

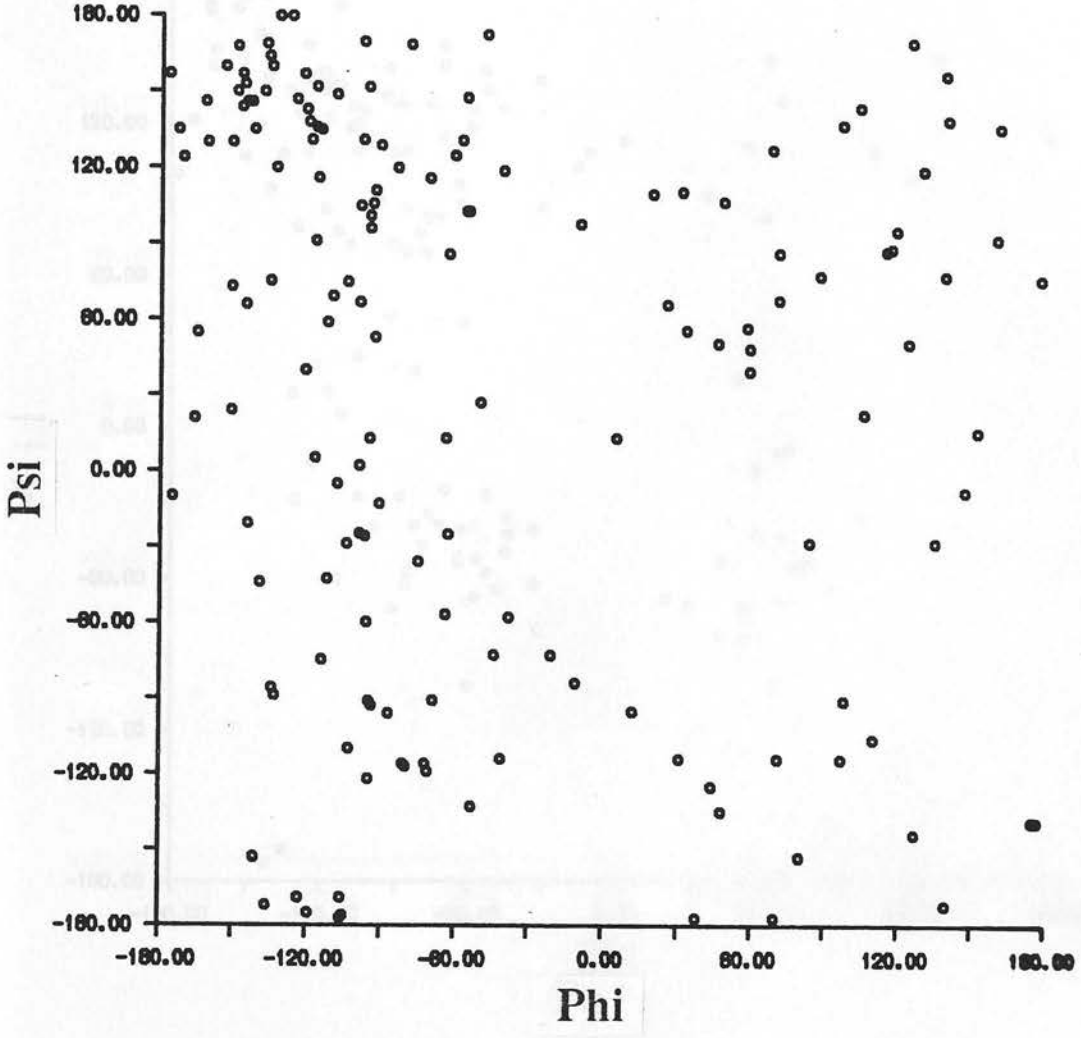


Figure 6.6 (a) Ramachandran plot of Blg lattice Y model at the beginning of this study.

Expected regions for protein polypeptide are:

	Phi (°)	Psi(°)
anti-parallel ß-sheet	-30 to -170	+40 to +170
alpha-helix	-60 to -120	-20 to -75

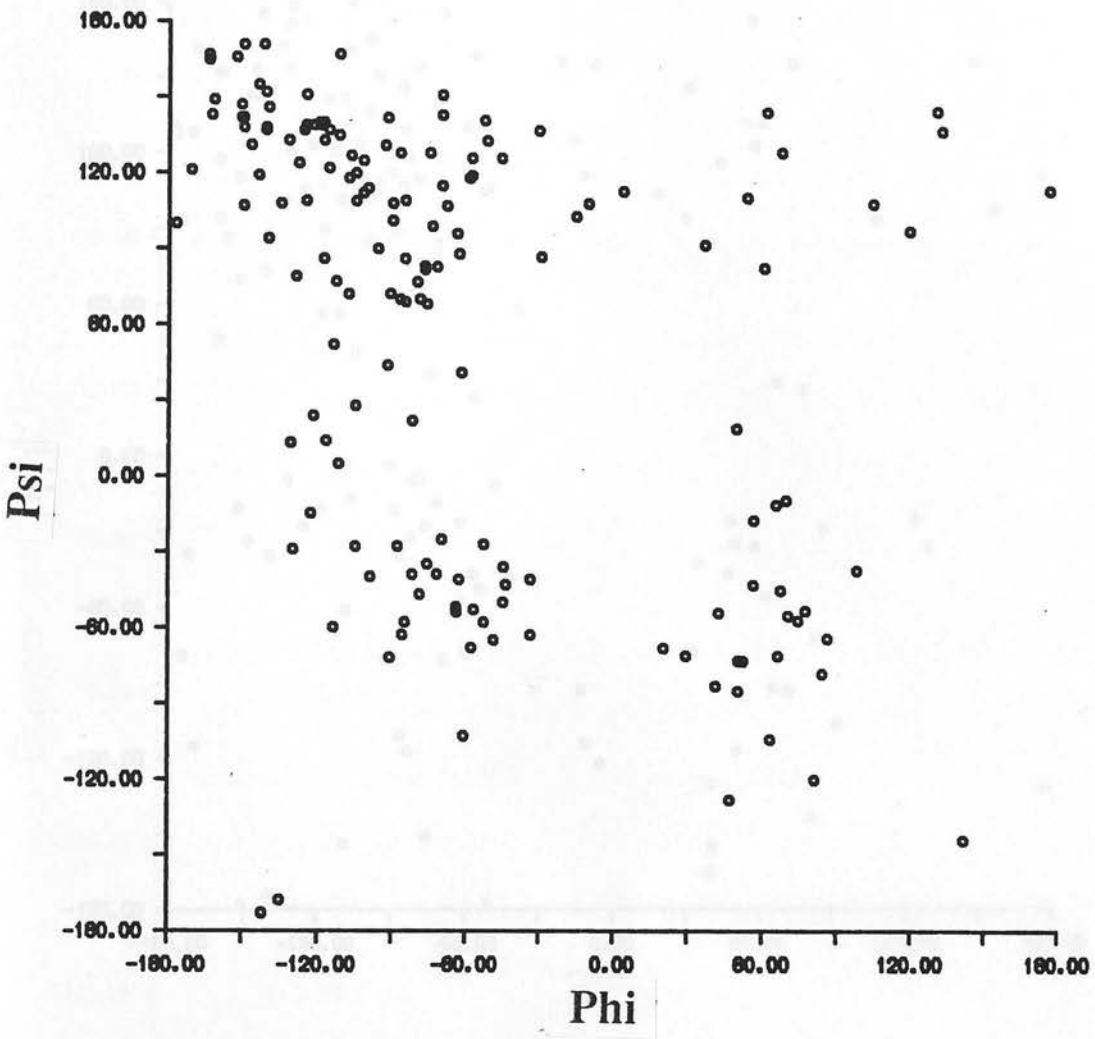


Figure 6.6 (b) Ramachandran plot of Blg lattice Y model after molecular dynamics refinement using MDXREF.

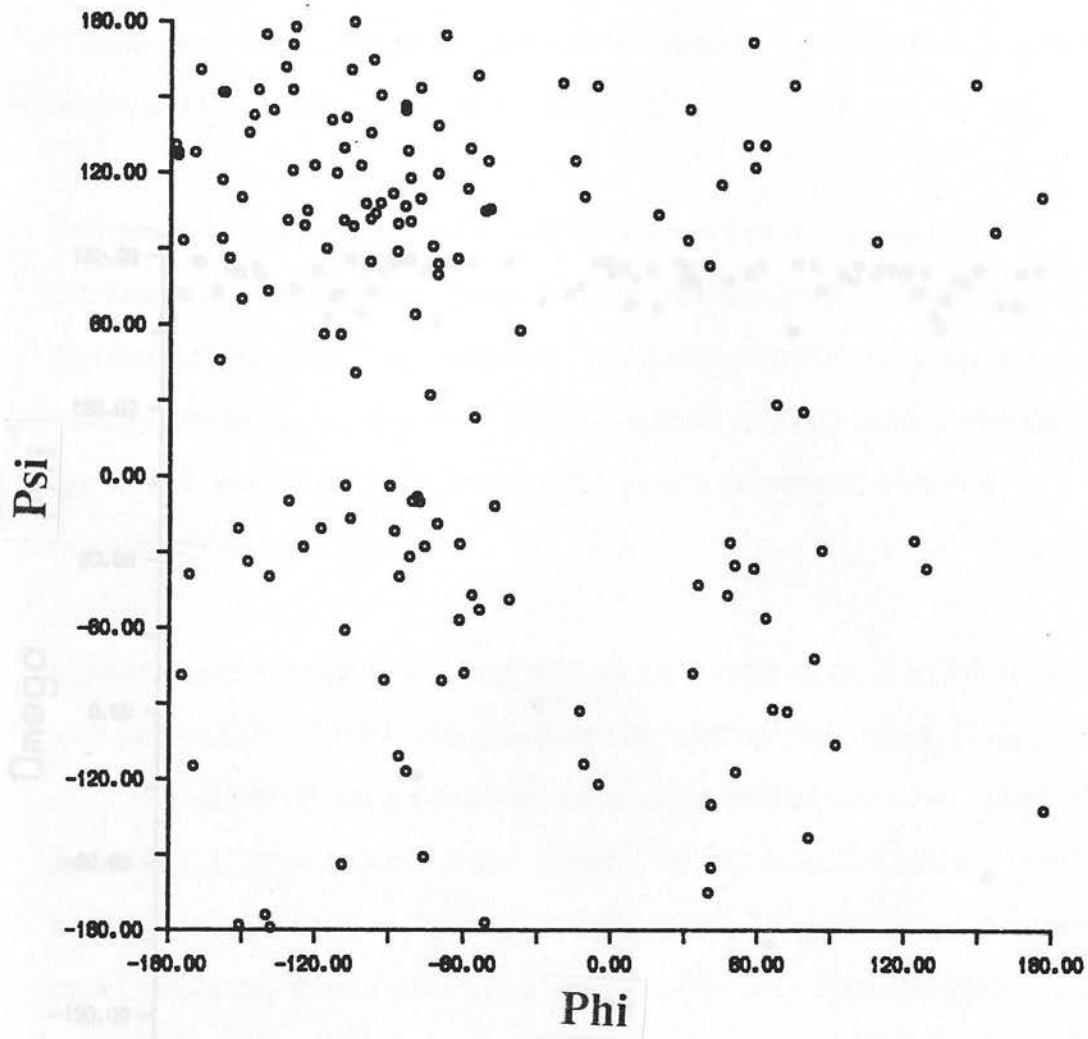


Figure 6.6 (c) Ramachandran plot of the final Blg lattice Y model.

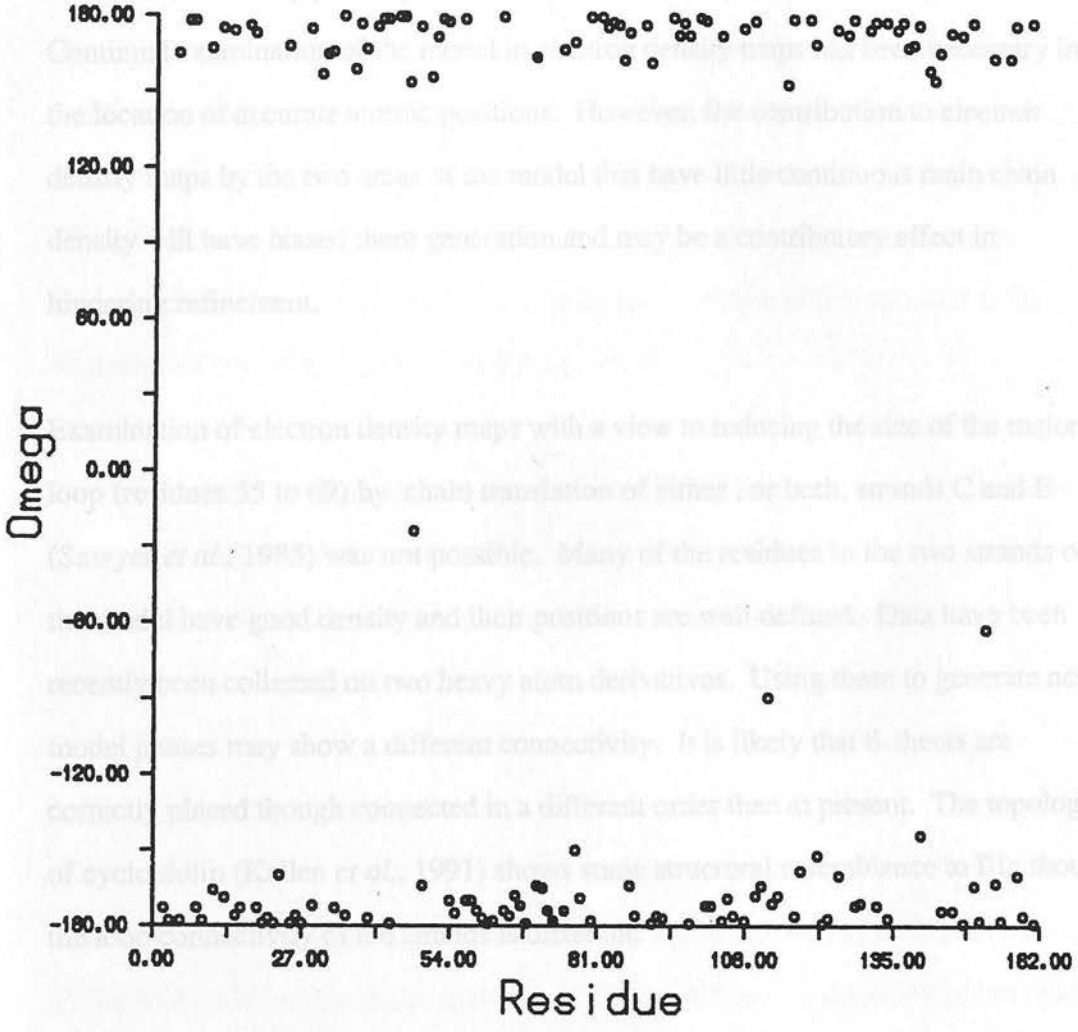


Figure 6.7 Omega plot of final Blg lattice Y model.

residues are prolines. Although cis-prolines are not uncommon (Stewart *et al.*, 1990) it seems likely that these residues have been introduced during the various modelling stages, and will need to be manually fixed.

Refinement has only partially led to the production of a more accurate structure. Continual examination of the model in electron density maps has been necessary in the location of accurate atomic positions. However, the contribution to electron density maps by the two areas of the model that have little continuous main chain density will have biased there generation and may be a contributory effect in hindering refinement.

Examination of electron density maps with a view to reducing the size of the major loop (residues 55 to 69) by chain translation of either , or both, strands C and E (Sawyer *et al.*, 1985) was not possible. Many of the residues in the two strands of the model have good density and their positions are well defined. Data have been recently been collected on two heavy atom derivatives. Using these to generate new model phases may show a different connectivity. It is likely that ß-sheets are correctly placed though connected in a different order than at present. The topology of cyclophilin (Kallen *et al.*, 1991) shows some structural resemblance to Blg though the loop connectivity of the strands is different.

Conversion of 20 residues to Ala, between pro-50 and lys-70, to reduce the contribution from this area towards the map production was attempted. The model was subjected to ten cycles of refinement in TNT and only succeeded in reducing the R-factor to 39%. Rebuilding the loop, to give accurate geometry, using consensus sequences from known structures resulted in a loop with better geometry but failed to reduce the R-factor below 38%. Neither of the techniques described here resulted in clear electron density around the area of interest.

The position of this loop is obviously going to be a problem in future refinement. It seems likely that within the crystal this loop is quite mobile and its position may never be completely resolved.

The use of data sets from different sources has undoubtedly been a contributory factor in holding up refinement. It is also possible that the problems encountered are a result of occasional crystals, although having apparently identical morphology, are of a different unit cell . This may explain the problems encountered when merging the different high resolution crystal sets together, which resulted in the omission of crystal C (see section 6.2.6).

The introduction of a new low resolution set collected from one crystal has already shown a model with a better R-factor than has been achieved before. Further refinement with high resolution data and a new set of phases should result in an accurate, high resolution structure.

6.3 Molecular Replacement.

The term molecular replacement was originally coined by Michael G. Rossmann ("The Molecular replacement method", 1972, and references therein) and has recently been reviewed (Rossmann, 1990). The technique is concerned with using non-crystallographic symmetry within the asymmetric unit to help phase the X-ray data from the unknown. Non-crystallographic symmetry results in the diffraction pattern containing redundant information. This symmetry can be expressed as

$$x_2 = [c]\, x_1 + d \qquad (5)$$

where [c] is a rotation matrix and d is a translational vector. This operation becomes crystallographic symmetry if it holds throughout the crystal rather than just within the region of molecules 1 and 2. The redundant information is used to solve the phase problem in the unknown. The solution of virus structures relies heavily upon non-crystallographic symmetry to improve the phasing of the structure. Initial phasing of the foot and mouth disease virus (FMDV) structure (Acharya *et al.*, 1989) was accomplished by obtaining initial phases from electron density from HRVI4 (Bernstein *et al.*, 1977) and phase extension was by molecular averaging and solvent flattening.

The realisation that many protein folds are similar, if not identical, has led to the use of molecular replacement as a tool for phasing unknown structures, almost "ab-initio". Finding the relative position of the unknown involves a search for similar/identical self Pattersons, but in a different orientation, using information derived from the unknown and the search structure. Here we use molecular replacement to utilise a known structure to help solve a related, and unknown structure.

The orientation of the blg lattice Y model within the cell of salt free grown crystals (see Chapter 5 for details) was examined using the program MERLOT (Fitzgerald, 1988).

6.3.1 The Molecular Replacement Program MERLOT.

MERLOT is an integrated package of programs which are designed to allow the user to best position/orientate a model structure within the cell of an unknown crystal. The molecular replacement follows three steps:

1. The determination of the Euler angles that describe the orientation of the known model in the unknown cell. (Euler angles can be described by a rotation α about the z axis, β about the new y axis, and finally γ about the new z axis. The rotation matrix can be seen in figure 6.8)

2. The determination of the translation of the properly orientated model.

3. Refinement of the above values.

The programs are arranged into 5 groups, plus a library, with each group sharing computer resources.

The flow diagram (Figure 6.9) describes the programs within the package involved in the search for the orientation of the model. After processing the coordinates, structure factors are calculated from the model in a cell with P1 symmetry. Both Patterson functions, i.e. for both model and unknown, are then generated in terms of spherical harmonics (a method which allows for quicker comparison) before a self rotation is completed. The self rotation allows the diffraction pattern to be searched for areas that contain similar information, therefore giving information on local symmetry of the unknown molecule. This is output as a polar angle map. Any rotation can be described by a spin about an appropriately chosen axis and the spherical polar angles, Phi and Psi, describe the longitude and coaltitude of this axis, Chi describes the rotation around it. Figure 6.8 shows the matrix describing polar angle rotations. A cross rotation is then carried out by the method of Crowther (1972) and outputs a series of peaks which can be analysed. The program ROTSYM calculates the symmetry mates of each peak, and checks to see if any of the peaks found are related.

Matrix in terms of Eulerian angles θ_1, θ_2 and θ_3

$-\sin\theta_1 \cos\theta_2 \sin\theta_3 + \cos\theta_1 \cos\theta_3$	$\cos\theta_1 \cos\theta_2 \sin\theta_3 + \sin\theta_1 \cos\theta_3$	$\sin\theta_2 \sin\theta_3$
$-\sin\theta_1 \cos\theta_2 \cos\theta_3 - \sin\theta_1 \cos\theta_3$	$\cos\theta_1 \cos\theta_2 \cos\theta_3 - \sin\theta_1 \sin\theta_3$	$\sin\theta_2 \cos\theta_3$
$\sin\theta_1 \sin\theta_2$	$-\cos\theta_1 \sin\theta_2$	$\cos\theta_2$

Matrix in terms of Polar angles ϕ and ψ and rotation angle χ

$\cos\chi + \sin2\psi \cos2\phi (1-\cos\chi)$	$\sin\psi \cos\psi \cos\phi (1-\cos\chi) - \sin\psi \sin\phi \sin\chi$	$-\sin\psi \cos\phi \sin\phi (1-\cos\chi) - \cos\psi \sin\chi$
$\sin\psi \cos\psi \cos\phi (1-\cos\chi) + \sin\psi \sin\phi \sin\chi$	$\cos\chi + \cos2\psi (1-\cos\chi)$	$-\sin\psi \cos\psi \sin\phi (1-\cos\chi) + \sin\psi \cos\phi \sin\chi$
$-\sin2\psi \sin\phi \cos\phi (1-\cos\chi) + \cos\psi \sin\chi$	$-\sin\psi \cos\psi \sin\phi (1-\cos\chi) \cos - \sin\psi \cos\phi \sin\chi$	$+ \sin2\psi \sin2\phi (1-\cos\chi)$

Figure 6.8 Eulerien and polar angle rotation matrices

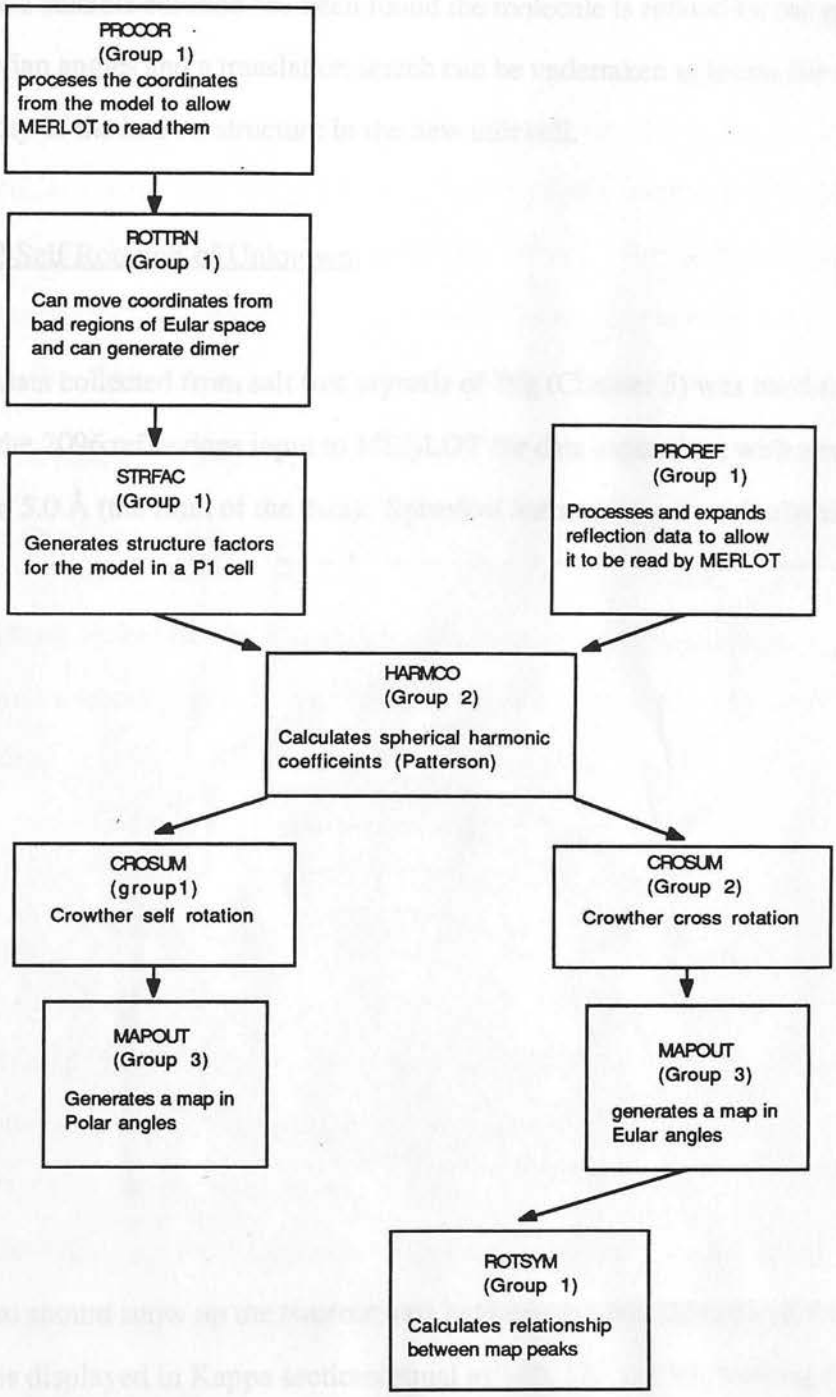
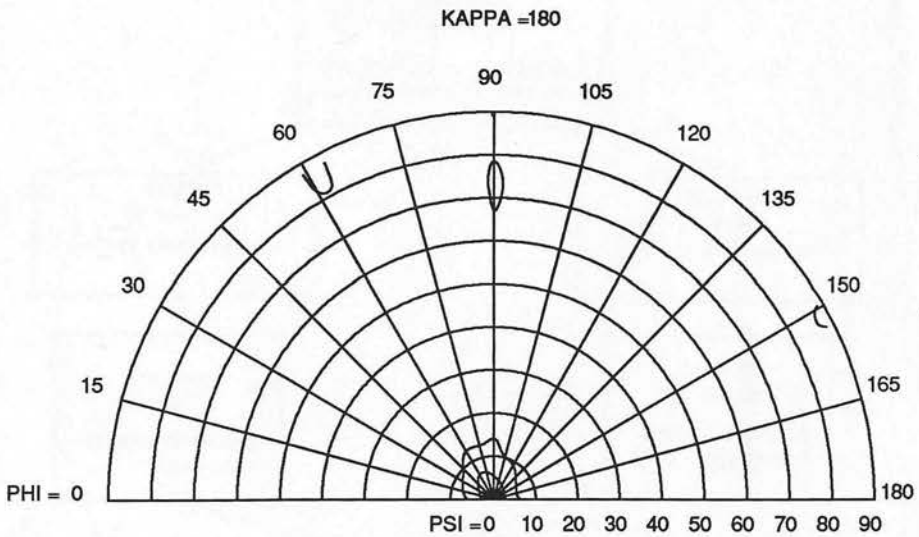


Figure 6.9 Flow diagram for program MERLOT. The diagram describes the steps taken in orientation of a known molecule in the unknown cell.

Figure 6.10 Self rotation map of Blg indicating twofold symmetry.



Once a suitable solution has been found the molecule is rotated by the appropriate Eulerian angles and a translation search can be undertaken to locate the centre of gravity of the known structure in the new unit cell.

6.3.2 Self Rotation of Unknown.

The data collected from salt free crystals of Blg (Chapter 5) was used as the unknown and the 2096 reflections input to MERLOT for data expansion, with a resolution cut off of 5.0 Å (the limit of the data). Spherical harmonics were calculated using a Patterson cut off radius of 29.9 Å. This value, which in theory should match the maximum molecular dimension but is not always possible, is coupled to the maximum resolution of the data being used. It is calculated from the following equation:

$$(2.0 \times \pi \times \text{ARAD}) / \text{RESMAX} < 37.6785 \qquad (6)$$

where ARAD is the cutoff value chosen, and RESMAX is the maximum resolution of the data being used. The value, therefore varied in different searches.

A polar angle map for the unknown was then calculated. A search with the Blg crystal should show up the twofold axis between the two subunits of the dimer. A map is displayed in Kappa sections equal to 180, 120 and 90, looking for twofolds, threefolds and fourfolds respectively. As expected a peak is seen to run through the section at 180 showing the presence of the dimer twofold in Blg. No fourfold is seen, indicating no dimer dimer relationship occurs.

6.3.3 Cross Rotation with Blg model.

The lattice Y model coordinates are read in protein data bank format. When reaching the point of translating the molecule, it is applied as a fraction of the unknown cell edges and, for this reason, the program is given cell information relating to the unknown crystal. A rotation of 180° around the z-axis generates the second molecule of the Blg dimer and allows for a search with the dimer if required. Bad regions of Euler space exist at β=0° and 180° since at these values, α and γ have similar and opposite effects respectively. One of the first solutions came out at α=45, β=5, γ=230. For this reason the search molecule was forced to α =50, β =100 and γ =75.

Lattice Y coordinates were then placed in a box of P1 symmetry, 100 Å on an edge. Structure factors are calculated for the model (be it monomer or dimer) to 4.54 Å, slightly higher than would ever be required. The Fcalc values are converted to spherical harmonics which are then used to perform a cross rotation between the model and the unknown. Using the monomer as a search molecule, and a search in the resolution range of 8 to 6 Å, an Euler map was generated with peak values as a fraction of the highest peak (which is set at 100%). The resulting map had 27 peaks over the recommended 65% cut off value. Above 80%, 9 peaks exist. Using the program ROTSYM with a Phi limit set at between 0 and 180 (the extent of our cross rotation polar map) the symmetry related peaks are generated and the relationships between them given.

Table 6.3 shows other searches with the number of peaks being produced on each occasion. The large number of peaks that are produced made it difficult to interpret the relationships found and to decide which of the peaks were the important ones to use as a solution.

Model	Model Limits	Data Limits	Search Limits	Max Peak Height (rms)	Ave. Peak Height (rms)	Min. Peak Height (rms)	Peaks > 80%	Peaks > 65%
monomer	4.8 Å	5.0 Å	9 - 5 Å	100 (4.9)	-1.17 (0)	-77 (-3.68)	9	27
dimer	4.8Å	5.0 Å	9 - 5 Å	100 (4.20)	0.67 (0)	-77.99 (-3.23)	12	74
dimer	4.54 Å	5.0 Å	8 - 5 Å	100 (4.25)	0.77 (0)	-83.14 (-3.59)	21	94
monomer	4.54 Å	5.0 Å	8 - 5 Å	100 (4.58)	0.00 (0)	-84.44 (-3.87)	11	69
monomer	5.0 Å	5.0 Å	8 - 6 Å	100 (4.78)	0.87 (0)	-59.06 (-2.89)	9	27
dimer	5.0 Å	5.0 Å	8 - 6 Å	100 (4.74)	0.20 (0)	-59.42 (-2.83)	13	84
monomer (bacbone)	5.0 Å	5.0Å	8 - 6 Å	100 (4.19)	0.69 (0)	-59.18 (-2.52)	30	-
dimer (backbone)	5.0 Å	5.0 Å	8 - 6 Å	100 (4.65)	0.41 (0)	-56.22 (-2.64)	7	25

Table 6.3 Summary of cross-rotation searches in MERLOT. Search molecule is the Blg lattice Y model. Unknown is from data collected from salt free Blg (see chapter 5)

It is possible that the nature of the protein i.e. a ß-barrel protein, causes many overlaps in different positions within the search and so many false peaks are generated which confuse the matter when trying to pick out unique and related peaks. Table 6.3 also indicates that a trial and error technique of changing the resolution limits (which also results in a different value for the Patterson cutoff radius) does not help in reducing the number of peaks that are produced. The problems that have been outlined earlier in this chapter about the lattice Y model, may mean that the search molecule is not accurate enough. This may contribute to noisy and uninterpretable peaks. Also, due to problems during the collection of the unknown data, reflections are quite weak and probably not as good as would be wished. This may also be a contributory factor in the failure to find good overlaps between the two data sets.

CHAPTER 7

Conclusions.

Chapter 7.

7.1 Introduction.

When the 2.8 Å Blg structure was first published (Sawyer *et al.*, 1985) a clue to the protein's function became apparent. Professor D. Phillips whilst refereeing the paper of the structure of human plasma retinol binding protein (RBP; Newcomer *et al.*, 1984) recognised the striking similarity in the overall fold of the two molecules.

This observation and the knowledge that a monomer of bovine Blg was capable of forming a tight 1:1 complex with retinol (Futterman & Heller, 1972; Hemley *et al.*, 1979; Fugate & Song, 1980), led to the hypothesis that Blg might be involved in the transport of retinol from milk to the newborn. A superposition of the backbones of RBP and Blg showed very close fit (Figure 7.1) and the characteristic cross-hatching effect of the ß-sheets.

Pervaiz & Brew (1985) while analysing homologous sequences proposed that the sequence similarity of the two proteins was consistent with the view that the two proteins would have similar three dimensional structures. This similarity was extended to protein HC (Lopez *et al.*, 1981).

Further sequence homologies, and two further crystal structures of related proteins, insecticyanin from tobacco hornworm (Holden *et al.*, 1987) and insect bilin-binding protein (Huber *et al.*, 1987), indicating the same fold, led to the identification of a superfamily of transport proteins. The proteins in the family, despite having low sequence similarity, are believed to have a very similar arrangement of secondary structural elements.

(a)

(b)

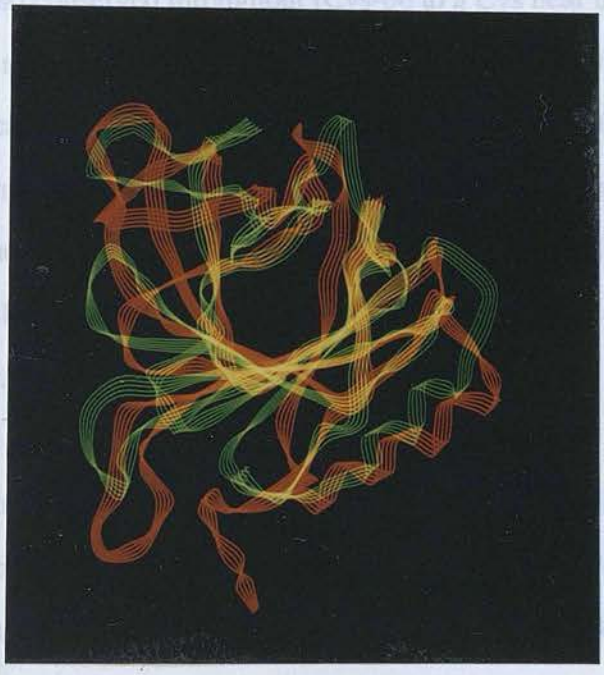


Figure 7.1 Superposition of RBP (red) and Blg (green). (a) shows the characteristic cross-hatching of the proteins and (b) shows the view looking into the RBP binding pocket.

There are some features in the primary sequence which appear to be invariant throughout the family. These were summarised by North (1989). The motifs are as follows:

- h - a - x - x - u - h - x - Gly - x - Trp - y - x - h - h -

(occurring near residue 20 of Blg)

- h - h - x - Thr - Asp - Tyr - x - x - y - h -

(occurring near residue 100 of Blg)

where a is usually acidic, u is often basic, y aromatic and h hydrophobic.

One disulphide is also conserved throughout (Cys-66 to a Cys near the C-terminus) as is a basic residue at position 124 (Blg numbering) and is normally Arg. Members of this superfamily can be seen in table X of appendix 2. Further evidence of the evolutionary relationship was given by Ali & Clark (1987) who showed the similarity in arrangement of the proteins' genes. Sawyer & Richardson (1991) have renamed this superfamily the lipocalycins owing to their ability to bind ligands within a hydrophobic pocket.

At the start of this study a preliminary model structure of Blg was available at 2.8 Å resolution which showed the overall fold of the molecule to be that of a ß-barrel. High resolution data had been collected to 1.8 Å resolution (Hambling, 1990) and extension of the present data should have given an accurate structure with defined side chain positions.

With this knowledge, the possibility of a high resolution structure, and the successful cloning of the ovine Blg gene in yeast (Paterson, 1991), the possibility exists to engineer

the protein to carry a specific ligand. The choice of ligand was that of a drug which causes problems when taken orally. As mentioned in Chapter 1, a molecule which could transport a compound through the stomach and ultimately allow its release in the intestines would be of considerable medical interest.

The refinement of the high resolution structure has not gone entirely to plan. The low to medium data that were available appeared sufficiently poor to hinder any extension of the model to a higher resolution. It is believed that this was a result of the data at low resolution which had been merged from a few different sets to give the existing data. Each set had been collected with a different instrument and, therefore, it is not surprising that each had different systematic errors which hindered a good combination.

Collection of a new and unique data set at 3.0 Å, the use of molecular dynamics and simulated annealing to refine the model into these new data proved to be more successful. Refinement by restrained least squares (using the package TNT) gives a model at 3.0 Å resolution with an R-factor of 21%. The model geometry is not yet as accurate as it could be, but future model building with inclusion of further, newly collected data to higher resolution should help in this matter. Table 7.1 lists data statistics and gives the mean deviation from the ideal values for the parameters of the final model.

This model can be used to choose residues as targets for genetic engineering. Residues within the hydrophobic core can be mutated such that the binding of the ligand of interest will be optimised. Lys-70 is one such suitable target, sitting at the mouth of the pocket where it may hinder the entry or exit of a prospective drug molecule (Paterson, 1991). Substitution by a smaller residue (e.g. Asn) will show whether the size or the charge is important in this regard.

Table 7.1 Statistics for final Blg model.

(Statistics taken from final model used with new data set at 3.0 Å)

Data Used in Refinement.

Starting Model	1279 protein atoms from Blg lattice Y
Program for model building	FRODO
Initial R-factor	47.8 % (model from Hambling, 1990)
Min. Resolution	10 Å
Max. resolution	3.0 Å
Structure amplitudes	$F_o \geq 1\sigma F_o$
Number of reflections	3032
Final R-factor	20.1 %

Deviation Statistics for different models.*

Model:	1	2	3	4	5
Mean bond deviation (Å)	0.429	0.071	0.461	-0.034	--0.073
Mean bond angle deviation (Å)	15.148	3.546	14.044	-0.037	1.265
Mean deviation for planar groups (Å)	-0.189	0.101	0.084	-0.040	-0.063
Mean torsion angle deviation (Å)	1.097	2.300	1.688	3.832	1.276
Mean trigonal atom non-planarity (Å)	-0.133	0.074	0.208	-0.019	-0.041
Bad contact distance, mean deviation (Å)	0.355	0.296	0.368	0.219	0.216
R-factor (%)	34.9	34.6	35.0	24.0	20.1

Models:

1 = before loop rebuilding

2 = after loop rebuilding

3 = model rebuilt before going into XPLOR

4 = after XPLOR

5 = final model

* If value is negative then mean value is below ideal value.

The model can be used as a basis to discuss the proteins' properties with respect to some of the residues.

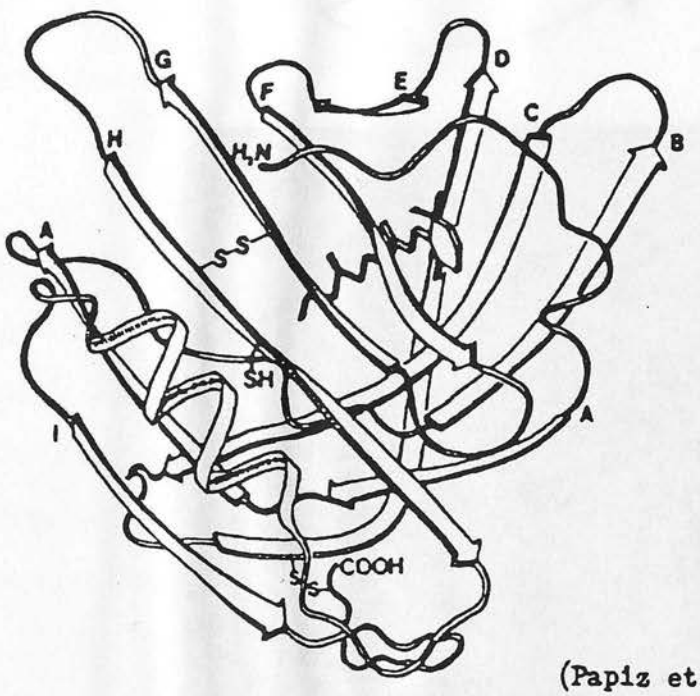
7.2 Structural analysis.

The structures of two small molecules served as a good introduction to the technique of X-ray crystallography. In the solving of the 8-OH DPAT structure, the importance of obtaining abundant and good quality crystals was outlined. Only one poor crystal was ever obtained which resulted in a structure which was not as accurate as EP092 (respective R-factors were 16.54% and 4.72%). It also served as a good introduction to using crystallographic programs and the use of molecular graphics for structural examination (Chapter 2).

Blg is predominantly made up of ß-sheet. Eight of the nine strands form anti-parallel sheets which wrap round to form the hydrophobic pocket. Figure 7.2 shows the overall fold and indicates the secondary structural elements. Strand A takes part in both upper and lower sheets of the barrel and interacts with strands F and G to close the barrel at its lower end. Strand I is involved in monomer-monomer association to give a dimer. Hydrophobic interactions between Ile-29 and Ile-147, and packing of the symmetry related His-146 is believed to hold the molecules together.

Blg contains one three turn α-helix which is amphipathic in nature (Figure 7.3) with the hydrophilic residues lying on the side exposed to solvent and hydrophobic residues lying inwards to the cleft defined by Monaco *et al.* (1987) as the retinol binding site.

```
        |--- A ---|          |- B -|   |-- C ---|
LIVTQTMKGL DIQKVAGTWY SLAMAASDIS LLDAQSAPLR VYVEELKPTP EGDLEILLQK
  |----- D ------|     |E|    |--- F ---|     |- G -|        |---
WENDECAQKK IIAEKTKIPA VFKIDALNEN KVLVLDTDYK KYLLFCMENS AEPEQSLVCQ
H -|     |-- HELIX --|    | I -|
CLVRTPEVDD EALEKFDKAL KALPMHIRLS FNPTQLEEQC HI
```

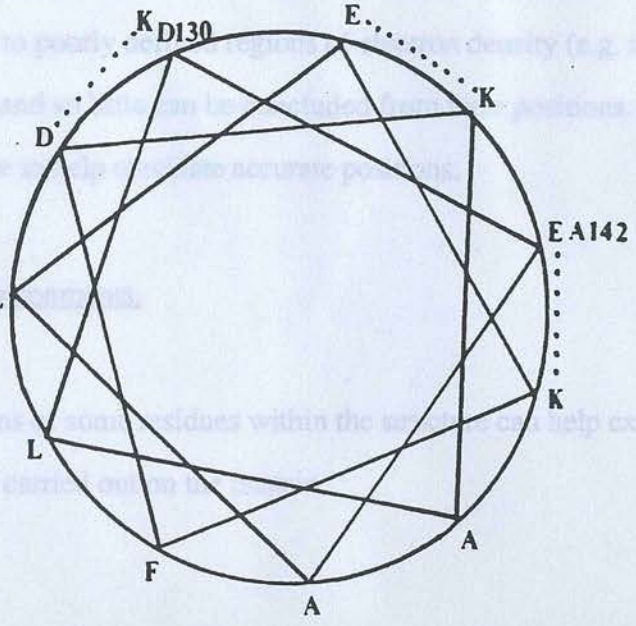


(Papiz et al.1986)

Figure 7.2 (a) Primary sequence of bovine Blg-A indicating secondary structure elements as indicated by Papiz *et al* (1986).

(b) Cartoon diagram showing overall fold of Blg lattice Y. Strands are labelled A to I

(a)



(b)

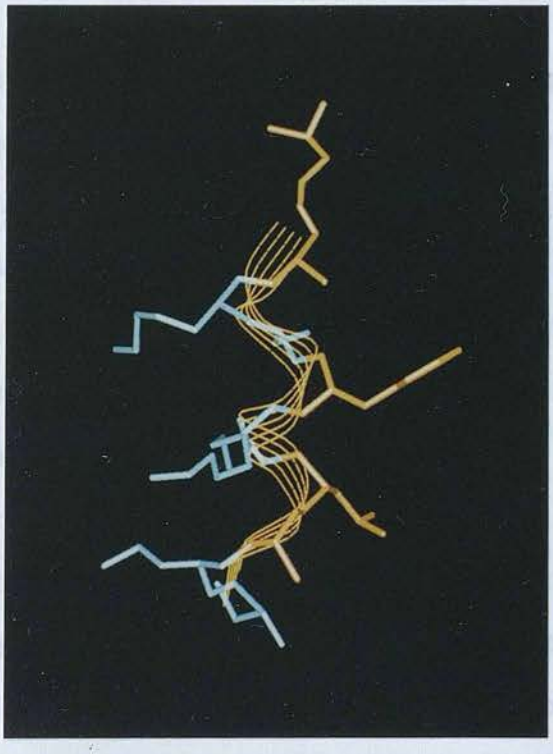


Figure 7.3 The amphipathic helix of Blg. (a) Diagramatic representation with broken lines indicating hydrogen bonds. (b) photograph showing hydrophilic residues (blue) and hydrophobic residues (orange)

The loop regions of the Blg model are the major uncertainty within the structure. Their flexibility lends to poorly defined regions of electron density (e.g. residues 55-70, see Chapter 6) and so little can be concluded from their positions. The use of template loops did little to help elucidate accurate positions.

7.2.1 Amino Acid Environments.

Examining the positions of some residues within the structure can help explain some of the solution studies carried out on the protein.

7.2.1.1 Tryptophan.

Blg contains two Trp residues, Trp-19 and Trp-61 (Figure 7.4). Trp-61 lies on the badly defined loop of the model but may form a crystal contact. Owing to the uncertainty of this loop no firm conclusions can be made. Trp-19 is more well defined and lies at the base of the hydrophobic pocket and is cut off from the solvent by Arg 124. Acylation studies correctly show that only 50% of the Trp residues in Blg are accessible and can be acetylated (Brown *et al.*, 1988). This is consistent with Trp-19 being inaccessible and Trp-61 being readily so.

North (1989) proposes that the position of Trp-19 may act as a switch mechanism to inform any receptor whether the protein has a ligand bound since the orientation of the residue in the X and Y lattice differs slightly. It seems more probable, however, that any interaction would be at the open end of the molecule. The Trp is located at the opposite end and so its function as such seems unlikely

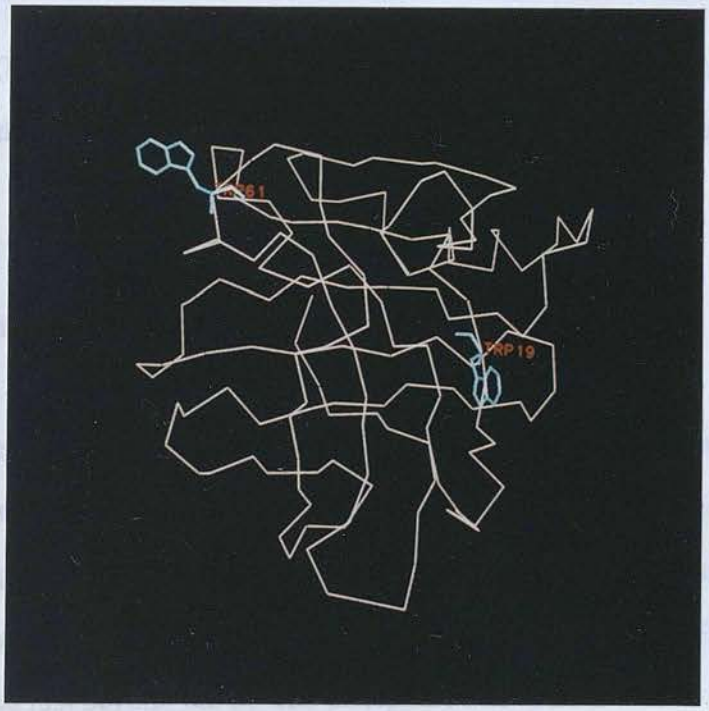


Figure 7.4 The locations of the tryptophan residues in Blg.

7.2.1.2 Tyrosine.

Four Tyr residues exist all of which appear to lie on one side of the monomer (Figure 7.5). Tyr -102 is buried in the cleft between the α-helix and strands G and H. Townend *et al.* (1969) indicate that only two can be acetylated. These are identified as Tyr-20 and Tyr-42, both of which are exposed. Tyr-99 is only partially accessible and may form an H-bond with Asp-11.

7.2.1.3 Phenylalanine.

Four Phe residues exist and the structure shows Phe-136, Phe-82 and Phe-105 all to be buried (Figure 7.6). Phe-136 lies in the cleft next to the α-helix and the other two lie within the hydrophobic pocket. Phe-151 is close to the C-terminus and is exposed to solvent.

7.2.1.4 Histidine.

His-146 lies on the I-strand and its stacking with its symmetry related mate contributes towards the dimerisation of the protein. Cleavage of the C-terminus by carboxypeptidase-A indicated that His-161 had to lie near the surface of the protein (Greenberg & Kalan, 1965). The structure shows that this is so (Figure 7.7) and His-161 lies just beyond the 66-160 disulphide interaction.

7.2.1.5 Lysine.

The 15 Lys residues are evenly distributed over the surface of the molecule and most

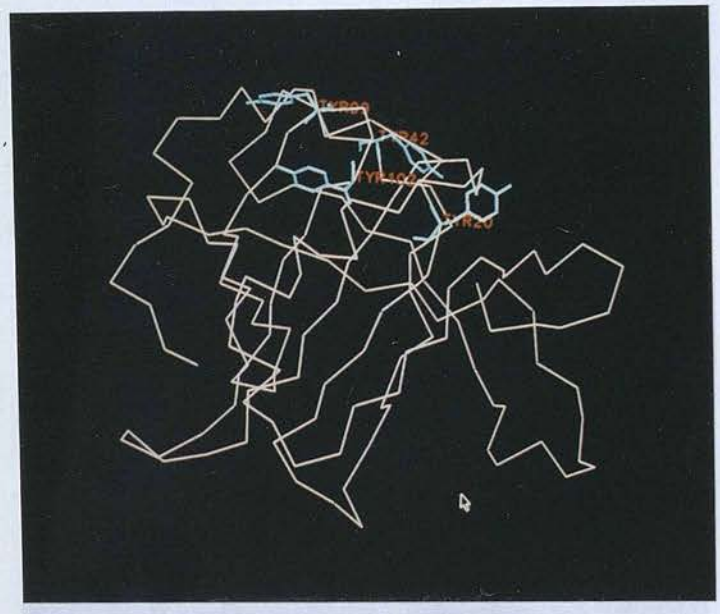


Figure 7.5 Location of the Tyrosine residues in Blg.

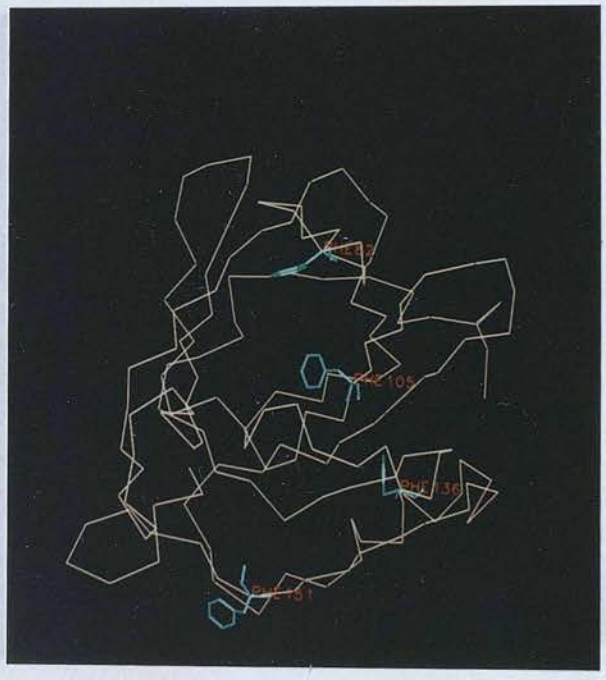


Figure 7.6 Location of the Phenylalanine residues in Blg

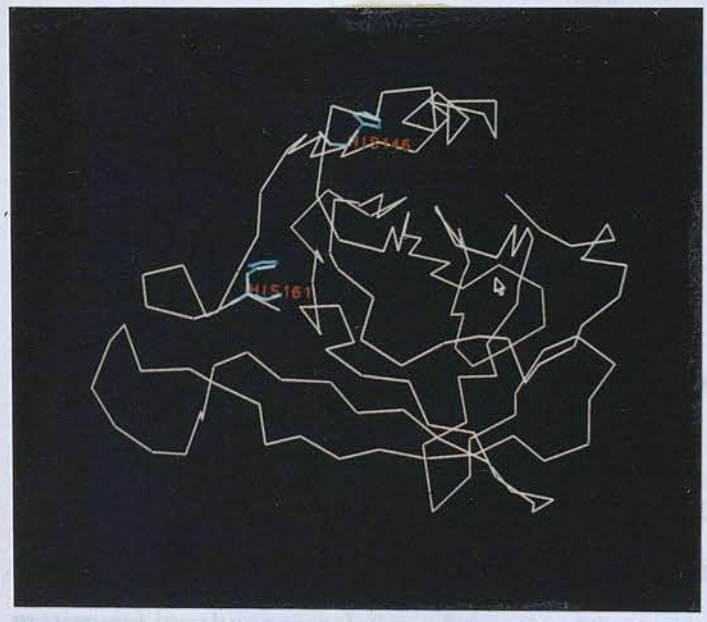


Figure 7.7 Location of the Histidine residues in Blg

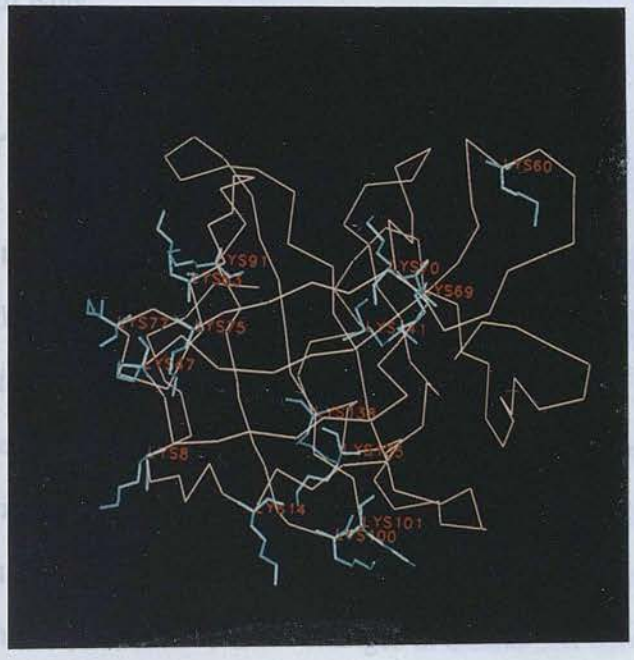


Figure 7.8 Location of the Lysine residues in Blg

look as though they are solvent accessible (Figure 7.8). Lys-135 and Lys-138 lie on the hydrophilic side of the α-helix and each is involved in a hydrogen bond with a neighbouring Glu (Figure 7.3a).

7.2.1.6 Cysteine.

Blg contains 5 Cys residues, 4 of which are involved in the two disulphides (Figure 7.9). In the lattice Y structure the free thiol is Cys-121. The sulphur atom of Cys-121 lies 10.91 Å from the sulphur on Cys-106 suggesting that any interchange of disulphide 106-119 to 106-121 would be unlikely without causing major structural movement involving translation of strands of ß-sheets. However, in pig Blg no Cys-119 exists and the corresponding disulphide occurs between Cys-106 and Cys121 (Conti *et al.*, 1986). The bond between Cys-66 and Cys-160 is conserved in all Blgs.

7.2.1.7 The Hydrophobic Cavity.

Strands A to H enclose the hydrophobic pocket which tapers from top to bottom. The pocket is lined by the following residues; Val-15, Trp-19, Pro-38, Val-41, Phe-82, Asp-88, Asn-90, Val-92, Val-94, Lys-103, Phe-105, Met-107, Asn-109 and Leu-122. The top of the pocket is lined by various polar residues and the opening is wide. Comparison of the lattice X and Y structures indicates that only minor changes have occurred to these residues, with 3 showing significant movement. Trp-19, Phe-105 and Phe82 show changes in orientation between the two structures and suggest that the position of any ligand in each structure may be slightly different (Yewdall, 1988).

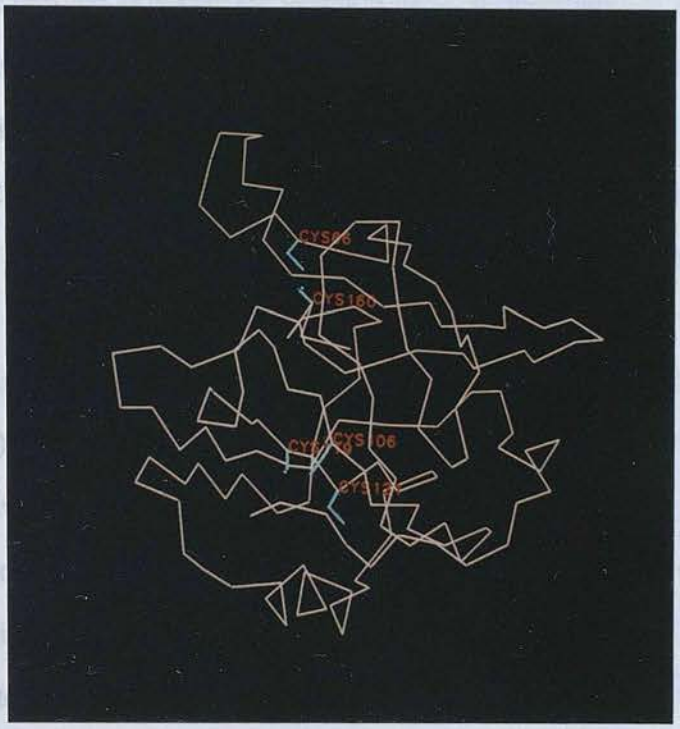


Figure 7.9 Location of cysteine residues in Blg. Cys-121 is the free thiol.

The possibility of engineering the protein now existed and the adequacy of Blg for such a purpose had to be examined.

7.3 Effects in the Stomach.

The protein is going to encounter extremely low pH within the stomach, a condition which it must survive to carry out its proposed function. Its resistance at low pH is borne out by the purification procedure (Aschaffenburg & Drewry, 1957) and has been investigated on a number of occasions (Timasheff & Townend, 1961; Townend *et al.*, 1960; McKenzie & Sawyer, 1967; Kella & Kinsella, 1988; Casal *et al.*, 1988). The conditions brought about by the stomach on food ingestion, i.e. exposure to acidic gut proteases, were examined in Chapter 3 and indicated that the protein will survive the action of these enzymes long enough to allow its passage to the intestines. Any ligand bound could be released here or internalised along with the protein. Marcon-Genty *et al.* indicate that 6 to 9 % of the protein crosses the gut wall intact

Hidalgo & Kinsella (1989) indicate that on the binding of a lipid (linoleic acid hydroperoxide) to Blg a tryptophan and a cysteine are oxidised. It seems likely that Trp-19 is involved since Trp-61 is located far from both the central cavity and the external channel (Figure 7.4). The oxidising of a Cys suggests that the binding site for the lipid is in the external cavity (see Figures 5.1 and 5.2). This leads to speculation as to whether the central cavity is too small to accommodate the ligand or if it is already filled. Binding of the lipid to Blg results in secondary products which react with amino groups to generate non-fluorescent dimers. The structure of these dimers is unknown. The binding of a ligand can, therefore, alter the structure of the protein. This may have resulted in the protein being more susceptible to gut proteases.

Griffen *et al.* (1989) studied the conformational stability of Blg by fluorescence spectroscopy. At pH 2.5 higher amounts of guanidine hydrochloride were required to denature the protein than at pH 6.9. This agreed with previous findings. The protein was, however, destabilised on the binding of retinol, requiring less guanidine hydrochloride than the native to allow denaturation, at both pH values.

In Chapter 3 the opposite is indicated under digestion conditions albeit at higher pH. The presence of retinol serves to stabilise the protein from the effects of trypsin. It should be noted that the dissociation constant of retinol from Blg is identical at both pH 7.5 and 2.0 (Fugate & Song, 1980) indicating that the binding would be just as tight at a low pH.

7.4 Allergic Reactions.

It would not be advantageous if a protein that was required to be taken orally to allow the introduction of a drug molecule caused allergic problems. Unfortunately Blg has been implicated as a causative agent in allergies that occur towards milk (Lebenthal, 1975). These allergies are generally towards children, and patients who would be expected to use the proposed drug compounds would normally be adults. Generally it is infants that are susceptible since antigen absorption is greater in the neonate due to the incomplete development of the gastrointestinal mucosal barrier (Udall *et al.*, 1981a,b). The rate of closure of the gut membrane to antigens is affected by the age and whether the infant is fed naturally or artificially.

The availability of monoclonal antibodies towards bovine Blg allowed further investigation into the regions responsible for the protein's allergenicity. It was hoped that a complex between the antigen binding fragment of the antibody and the Blg antigen could be obtained and crystallised. A structure of the complex would confirm

the protein's putative antigenic regions and possibly pick out flexible loops of the protein. The information gleaned may be useful and interesting from a pediatric point of view.

The first batch of monoclonal antibodies available proved to be inactive owing to the way they had been prepared before transportation from the USA. The production of polyclonal antibodies, which were purified by affinity chromatography, was undertaken to examine the techniques that were being used and proved their validity. Recently, a new batch of monoclonal antibodies has been received as an ammonium sulphate precipitate. The activity of these antibodies has been shown to be satisfactory (Chapter 4) and will warrant further investigation which will result in a complex that can be crystallised to yield a structure that will give information as to the important regions of the protein.

The program ELLIPSE (kindly supplied by Janet Thornton) creates an ellipsoid around a protein molecule and calculates a protrusion index for portions of the molecule. Blg was examined with the program and several areas were shown to be more exposed than the rest of the protein. ELLIPSE picks out four regions in particular. The first three sets of residues that are exposed are 41-47, 55-65 and 108-114, all of which stretch between ß-sheets as loop regions. A region prior to the α-helix, 124-130, is also chosen. None of the regions is spatially close to any other and none of them would be occluded by dimerisation. As a result of these regions being loops or turns it is reasonable to assume that some mobility would be associated with them. It is not clear, however, if this would lend to the assumption that these regions could be chosen as possible epitopes for the protein. Two of the chosen sites can be included in the antigenic region located by S-carboxymethylation (Otani *et al.*, 1985).

Smith-Gill *et al.* (1982) indicate that an epitope can occur over a wide area and encompass a range of secondary structural elements. Sheriff *et al.* (1987) showed that three distinct oligopeptide segments are involved in the lysozyme epitope. Two of these segments involve turn or loop regions, with the majority of the interaction with sheet structure.

However, an alternative possibility is that these mobile regions in Blg are responsible for immunogenically stimulating ß-lymphocytes and need not be involved in the antigenic determinant. The answer to this question must await the structure of a complex.

7.5 Crystallisation of Blg.

Because of the need to crystallise the low pH form of the protein, the conditions of crystallisation were investigated. Crystals were obtained at pH 3.1 (sp. grp. $P\,6_3$ a = b = 68.49 Å, c = 143.17 Å) from ammonium sulphate and are relevant to the main aims of this project. Data collected from this crystal form have still to be processed but should give relevant information on the protein structure at this pH and may give an indication why the protein is so stable at such a low pH. Acidic residues within the acid proteases probably play a role in their stability.

Data from crystals at pH 3.0, 5.2 (sp. grp. $P\,2_1$ a = 72.2 Å, b = 67.9 Å, c = 36.2 Å, ß = 92.0°), 6.8 (Yewdall, 1988) and 7.8 (Chapter 6) are now available and should help give an overall view of the protein through the physiological pH range. It may be possible to use molecular replacement to determine phases for the unknown structures. The most obvious choice of starting model would be the lattice Y Blg structure, but alternatively that of RBP (Cowan *et al.*, 1991) or insecticyanin (Holden

et al., 1987) could be used. Chapter 6 describes preliminary investigation into the orientation of the lattice Y model in the cell of the crystal form grown at pH 5.2. Too many peaks were produced and it was difficult to distinguish any result from the noise. It is possible that the ß-barrel nature of the protein results in many overlaps producing additional and false peaks. It may be, therefore, that these crystal forms will have to be solved by traditional isomorphous replacement methods. Work into the production of more crystals of higher quality is in progress.

Chapter 6 also indicates the importance of obtaining good crystals and a reliable source of protein. If crystals are good then accurate data can be collected and should result in less problems when it comes to processing and using the data at a later stage.

In summary, a low pH crystal form of bovine Blg exists and will allow a structure to be solved at pH 3.1. This will help to explain the resistance that the protein shows to the acid proteases. The hydrophobic pocket can be further examined and suitable residues chosen to mutate in the ovine gene with the help of a low pH structure and the high resolution lattice Y structure. Both crystals will facilitate the binding of a specific drug molecule. Examination of the protein's antigenicity will be helped once a complex of the active monoclonal antibody is crystallised and a structure solved. This is relevant to pediatrics and to the consumption of the protein by patients who are required to take the drug for whose delivery Blg has been tailored.

7.6 Function.

Blg is a protein about which much is known, though one property still remains elusive: its function. Various pieces of evidence seem to point towards the protein being involved in the transportation of a ligand to the suckling animal. Both retinol (Garrick & Williamson, 1986) and free fatty acid (Perez *et al.*, 1989) have been

shown to be associated with Blg in milk and so both represent possibilities.

It seems likely that whatever the complex, and however it is formed, i.e. whether in the hydrophobic pocket or in the exterior channel, the complex is transferred from mother to young via suckling. The complex survives intact through the stomach (see Chapter 3) and as it reaches the intestines the rise in pH brings about the Tanford transition. This may allow release of the ligand or the conformational change may result in the protein binding to a specific receptor which would then allow uptake of the ligand. Retinol uptake has been shown to increase in the presence of Blg (Said *et al.*, 1989) and specific receptors have been identified in the intestine of the neonate calf (Papiz *et al.*, 1986). The protein may then be broken down and used for nutrition.

It may also be possible that Blg has non-specific transport function in that it responsible for transporting of a variety of ligands from milk to the young. It may transport the molecule across the stomach and facilitate its uptake across the membrane in the intestines.

The requirement for the presence of Blg being linked to placental type has been disproved (appendix 1) and so its presence in some animals but not others remains a mystery. The fact that it is a member of the lipocalycin superfamily and that it has the ability to bind many ligands suggests that its function is, or was, one of transport. However, the ligand that it transports has to be correctly identified.

The successful cloning of the gene in *E. coli* and yeast, and the potential of further structures of the protein at different conditions of pH will lead to a more accurate hypothesis for the protein's function.

BIBLIOGRAPHY

Bibliography.

Acharya, R., Fry, E., Stuart, D., Fox, G., Rowlands, D. & Brown, F. (1989) The three-dimensional structure of foot-and-mouth disease virus at 2.9 A resolution. Nature **337** 709-716.

Akroyd, P. (1965) Acrylamide - gel electrophoresis of ß–lactoglobulins stored in solutions at pH 8.7. Nature **208** 488-489.

Ali, S. & Clark, A.J. (1988) Characterisation of the gene encoding ovine ß-lactoglobulin. Similarity to the genes for retinol binding protein and other secretory proteins. J. Mol. Biol. **199** 415-426.

Allen, F.H., Kennard, O., Watson, D.G., Brammer, L., Orpen, A.G. & Taylor, R. (1987) J. Chem. Soc. Perkin Trans. **2** s1-s19.*

Amit, A.G., Mariuzza, R.A., Phillips, S.E.V. & Poljak, R.J. (1986) Three dimensional structure of an antigen-antibody complex at 2.8 A resolution. Science **233** 747-753.

Armstrong, J.McD., McKenzie, H.A. & Sawyer, W.H. (1967) On the fractionation of β–lactoglobulin and α–lactalbumin. Biochim. Biophys. Acta **147** 60-72.

Aschaffenburg, R. & Drewry, J. (1957) Improved method for the preparation of the crystalline β–lactoglobulin and α–lactalbumin from cow's milk. Biochem. J. **65** 273-277.

Aschaffenburg, R., Green, D.W. & Simmons, R.M. (1965) Crystal forms of β–lactoglobulin. J. Mol. Biol. **13** 194-201.

Axelsson, I., Jakobsson, I., Lindberg, T. & Benediktsson, B. (1986) Bovine β–lactoglobulin in human milk. Acta Pediatr. Scand. **75** 702-707.

Bain, J.A. & Deutsch, H.F. (1948) Studies on lactoglobulins. Arch. Biochem. **16** 221-229.

Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Bruce, M.D., Rodgers, J.R., Kennard, O., Shimanouchai, T. & Tasumi, M. (1977) The protein data bank: A computer-based archival file for macromolecular structures. J. Mol. Biol. **112** 535-542.

Blundell, T. & Johnson, L. (1976) Protein Crystallography. (Ed. B. Horecker, N. Kaplan, J. Marmur & H.A. Scheraga) Academic Press, London.

Bocelli, G., Catellani, M. & Chiusoli, G.P. (1984) Palladium-catalyzed carbon-carbon bond formation involving aromatic carbon-hydrogen activation. III Aspects of aromatic substitution and structure of 1-bromo-3-[3-(2-methylenecyclo hex-5-en--1-yl)bicyclo[2.2.1]hept-2-yl]-4-bicyclo[2.2.1]hept-2-yl benzene. J. Organomet. Chem. **279** 225-232.

*Tables of bond lengths determined by x-ray and neutron diffraction. Part 1. Bond lengths in organic compounds.

Bolognesi, M., Liberatori, J., Oberti, R. & Ungaretti, L. (1979) Prelimenary crystallographic data on buffalo β–lactoglobulin. J. Mol. Biol. **131** 411-413.

Brignon, G., Chtourou, A. & Ribadeau-dumas, B. (1985) Does β–lactoglobulin occur in human milk? J. Dairy Res. **52** 249-254.

Brock, J.H., Arzabe, F., Lampreave, F. & Pineiro, A. (1976) The effect of trypsin on bovine transferrin and lactoferrin. Biochim. Biophys. Acta **446** 214-225.

Brown, W.R., Newcomb, R.W. & Ishizaka, K. (1970) Proteolytic degradation of exocrine and serum immunoglobulins. J. Clin. Invest. **49** 1374-1380.

Brown, E.M., Pfeffer, P.E., Kumosinski, T.F. & Greenberg, R. (1988) Accessibility and mobility of lysine residues in ß-lactoglobulin. Biochemistry **27** 5601-5610.

Bruck, C., Portelle, D., Gilneur, C. & Bollen, A. (1982) One-step purification of mouse monoclonal antibodies from ascitic fluid by DEAE Affi-gel blue chromotography. J. Immunol. Meth. **53** 313-319.

Brunger, A.T., Kuriyan, J. & Karplus, M. (1987) Crystallographic R-factor refinement by molecular dynamics. Science **235** 458-460.

Casal, H.L., Kohler, U. & Mantsch, H.N. (1988) Structural and conformational changes of ß-lactoglobulin B: an infrared spectroscopic study of the effect of pH and temperature. Biochim. Biophys. Acta **957** 11-20.

Chobert, J.M., Dalgalarrondo, M., Dufour, E., Bertrand-Harb, C. & Heartle, T. (1991) Influence of pH on the structural changes of β–lactoglobulin studied by tryptic hydrolysis. Biochim. Biophys. Acta **1077** 31-34.

Cogan, U., Kopelman, M., Mokady, S. & Shinitzky, M. (1976) Binding affinities of retinol and related compounds to retinol binding proteins. Eur. J. Biochem. **65** 71-78.

Colman, P.M., Laver, W.G., Varghose, J.N., Baker, A.T., Tulloch, P.A., Air, G.M. & Webster R.G. (1987) Three-dimensional structure of a complex of antibody with influenza virus neuraminidase. Nature **326** 368-363.

Conti, A. & Godovac-Zimmermann, J. (1990) Antibodies raised against bovine β–lactoglobulin react with B2-microglobulin. Biol. Chem. Hoppe-Seyler **371** 261-263.

Cowan, S.W., Newcomer, M.E. & Jones, T.A. (1990) Crystallographic refinement of human serum retinol binding protein at 2.0 Å resolution. Proteins, Structure, Function and Genetics **8** 44-61.

Creamer, L.K., Paryy, D.A.D. & Malcolm, G.M. (1983) Secondary structure of bovine ß-lactoglobulin B. Arch. Biochem. Biophys. **227** 98-105.

Conti, A., Godovac-Zimmermann, J.,Pischner, F., Liberatori, J. & Braunitzer, G. (1986) Pig ß-lactoglobulin I (*Sus scrofa domestica*, Artiodactyla). Biol. Chem. Hoppe-Seyler **367** 871-878.

Crowfoot, D.M. & Riley, D.P. (1938) An X-ray study of Palmers lactoglobulin. Nature **141** 521-522.

Crowther, R.A. (1972) In "The molecular replacement method" The fast rotation function. Ed. Rossmann, M.G., Gordon & Breach.

Davies, D.R. & Rich, A. (1971) Protein crystallisation: microtechniques involving vapor diffusion. In Mets. Enzymol. (W.B. Jacoby, Ed.) **22** Acad. press, NY. p.266.

Davies, D.T. (1974) The quantitative partition of the albumin fraction of milk serum proteins by gel chromotography. J. Dairy Res. **41** 217-228.

Dodin, G. Andrieux, M. & Al Kabbani, H. (1990) Binding of ellipticine to ß-lactoglobulin. Eur. J. Biochem. **193** 697-700.

Fitzgerald, P.M.D. (1988) MERLOT - Integrated package of computer programs for the determination of crystal structures by molecular replacement. Merck Sharp & Dohme, New Jersey, USA.

Fortier, S., Erman, M.G., Langs, D.A. & DeTitta, G.T. (1980) The structure of thromboxane B2. Acta Cryst. **B36** 1099-1013.

Fox, G.C. & Holmes, K.C. (1966) An alternative method to solving the layer scaling equations of Hamilton, Rollett and Sparks. Acta Cryst. **20** 886-891.

Frushour, B.G. & Koenig, L.J. (1975) Raman studies of the crystalline, solution and alkaline denatured states of ß-lactoglobulin. Biopolymers **14** 649-662.

Fugate, R.D. & Song, P.S. (1980) Spectroscopic characterisation of β–lactoglobulin - retinol complex. Biochim. Biophys. Acta **625** 28-42.

Fujinaga, M., Gros, P. & van Gunstern, W.F. (1989) Testing the method of crystallographic refinement using molecular dynamics. J. Appl. Cryst. **22** 1-8.

Futterman, S. & Heller, J. (1972) The enhancement of fluorescence and the decreased susceptibility to enzyme oxidation of retinol complexed with bovine serum albumin, ß-lactoglobulin and the retinol binding protein of human plasma. J. biol. Chem. **2478** 5168-5172.

Galloway, T.S. & van Heyningen, S. (1987) Binding of NAD by cholera toxin. Biochem. J. **244** 225-230.

Garavito, M. & Rosenbusch, J.P. (1980) Three dimensional crystals of an integral membrane protein: an initial X-ray analysis. J. Cell Biol. **86** 327-329.

Garrick, J.D. & Williamson, F.W. (1986) Is the role of ß-lactoglobulin to serve as a retinol carrier in milk? Bsc. Hons. project thesis, University of Aberdeen.

Giesecke, J. (1980) The crystal structure of (+)-2-Dipropylamino-5- hydroxy tetralin

hydrochloride. Acta Cryst. **B36** 110-114.

Goldman, P., Peppercorn, M.A. & Goldin, B.R. (1974) Drugs metabolised by intestinal microflora. In "Drug interactions" (Eds, P.L. Monselli, S. Garattini, S.N. Cohen) Raven press, NY. pp. 91-102.

Gould, R.O. & Taylor, P. (1983) CALC. Interactive program for molecular geometry. University of Edinburgh, Scotland.

Green, D.W., Ashaffenburg, R., Camerman, A., Coppola, J.C., Diamond, R.D., Dunnill, P., Simmons, R.M., Komorowski, E.S., Sawyer, L., Turner, E.M.C. & Woods, E.F. (1979) Structure of bovine β–lactoglobulin at 6 A resolution. J. Mol. Biol. **131** 375-397.

Greenberg, R. & Kalan, E.B. (1965) Studies on ß-lactoglobulins A, B and C. II. Preparation of modified proteins by treatment with carboxypeptidase-A. Biochem. **4** 1660-1666.

Griffen, W.G., Griffen, M.C.A. & Martin S.R. (1989) Conformational studies of the ß-lactoglobulin/retinol complex. Presented at the 32nd IUPAC congress, Stockholm.

Gros, P., Fujinaga, M., Mattevi, A., Velliex, F.M.D., van Gunsteren, W.F. & Hol, W.G.F. (1989) Protein structure refinement by molecular dynamics techniques. Proc. Daresbury study weekend 1-15.

Groves, M.L., Hipp, N.J. & McMeekin, T.L. (1951) Effect of pH on the denaturation of ß-lactoglobulin and its dodecyl sulphate derivative. J. Am. chem. Soc. **72** 2790-2793.

Gu, W. & Brady, J.W. (1991) Molecular dynamics simulations of the whey protein ß-lactoglobulin. In Press.

Halliday, J.A., Bell, K. & Shaw, D.C. (1991) The complete amino-acid sequence of feline ß-lactoglobulin II and a partial revision of the equine ß-lactoglobulin II. Biochim. Biophys. Acta. **1077** 25-30.

Hambling, S.G. (1990) Crystallographic and solution studies of bovine β–lactoglobulin. Ph.D. thesis, University of Edinburgh.

Hambling, S.G., McAlpine, A.S. & Sawyer, L. (1991) β–lactoglobulin. Chapter in Milk Proteins, in Press. (Ed Fox, P., Elsevier publications).

Hambling, S.G., Yewdall, S.J., north, A.C.T. & Sawyer, L. (1987) Comparison of two crystal forms of ß-lactoglobulin. Acta Cryst. **A43** C-26.

Harding, M.M. (1990) Protein structural aspects of biotechnology, Comett course manual, Universities of Edinburgh & Nijmegen. 51-58.

Hemley, R., Kohler, B.E. & Siviski, P. (1979) Absorption spectra for the

complexes formed from vitamin-A and β–lactoglobulin. Biophys. J. **28** 447-455.

Hidalgo, F.J. & Kinsella, J.E. (1989) Changes induced in ß-lactoglobulin B following interactions with linoleic acid 1B-hydroperoxide. J. Agric. Food Chem. **37** 860-865.

Holden, H.M., Rypiewski, W.R., Law, J.H. & Rayment, I. (1987) The molecular structure of insecticyanin from the tobacco hornworm (*Manduca sexta L.*) at 2.6 Å resolution. EMBO J. **6** 1565-1570.

Huang, Q., Coleman, J.W. & Stanworth, D.R. (1985) Investigation of the allergenicity of β–lactoglobulin and its cleavage fragments. Int. Archs. Allergy Appl. Immunol. **78** 337-344.

Huber, R., Schnieder, M., Epp, O., Mayr, I., Messerchmitt, A., Pflugrath, J. & Kayser, H. (1987) Molecular structure of the bilin binding protein (BBP) from *Pieris brassicae* after refinement at 2.0 Å resolution. J. Mol. Biol. **195** 423-513.

International Tables for X-ray Crystallography (1974) **Vol. IV.** Birmingham: Kynoch Press (Present distributor Kluwer Academic Publishers, Dordrecht).

Jacobs, G. & Cunningham, L.W. (1968) Creatine kinase. The relationship of trypsin susceptibility to substrate binding. Biochemistry **7** 143-151.

Jakobsson, I., Lindberg, T., Benediktsson, B. & Hansson, B.G. (1985) Dietary bovine β–lactoglobulin is transferred to human milk. Acta Pediatr. Scand. **74** 342-345.

Johansson, A.M., Nilsson, J.L.G., Karlen, A., Hacksell, U., Svensson, K., Carlsson, A., Kenne, L. & Sundell, S. (1987) C3-methylated 5-hydroxy-2-(dipropylamino) tetralins: Conformational and steric parameters of importance for central dopamine receptor activation. J. Med. Chem. **30** 1135-1144

Johnson, C.K. (1965) ORTEP. Report ORNC-3794. Oak Ridge National Laboratory, Tennessee, USA.

Johnson, J. (1989) Proteins grown in space reveal their detailed structure. New Scientist 16th Dec, **124** 21 only.

Jones, T.A. (1978) A graphics model building and refinement system for macromolecules. J. Appl. Cryst. **11** 268-272.

Kallen, J., Spitzfaden, C., Zurini, M.G.M., Wider, G., Widmer, H. Wüthrich, K. & Walkinshaw, M.D. (1991) Structure of human cyclophilin and its binding site for cyclosporin A determined by X-ray crystallography and NMR spectroscopy. Nature **353** 276-279.

Karlsson, A., Petterson, C., Sundell, S., Arvidsson, L.E. & Hacksell, U. (1988) Improved preparation, chromatographic separation, and x-ray crystallographic determination of the absolute configuration of the enantiomers of 8-hydroxy-2-(dipropylamino)tetralin (8-OH DPAT). Acta Chemica Scand. **B42** 231-236.

Kato, A., Komatsu, K., Fujimoto, K. & Kobayashi, K. (1985) Relationship between surface functional properties and flexibility of proteins detected by the protease susceptibility. J. Agric. Food Chem. **33** 931-934.

Kella, N.K.D. & Kinsella, J.E. (1988) Enhanced thermodynamic stability of β–lactoglobulin at low pH. Biochem. J. **255** 113-118.

Kessler, E. & Brew, K. (1970) The whey proteins of pig's milk, isolation and characterisation of a ß-lactoglobulin. Biochim. biophys. Acta **200** 449-458.

Kirby, B.J., McAlpine, A.S., Sawyer, L., Taylor, P. & Blake, A.J. (1991) Structures of two 5-Hydroxy tryptamine receptor agonists. In Press.

Kohler, G. & Milstein, C. (1975) Continuous cultures of fused cells secreting antibody of predetermined specificity. Nature **256** 495-497.

Kuzmanoff, K.M. & Beattie, C.W. (1991) Isolation of monoclonal antibodies monospecific for bovine β–lactoglobulin. In Press.

Laemlli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **277** 680-685.

Lerner, R.A. (1982) Tapping the immunological repertoire to produce antibodies of predetermined specificity. Nature **299** 592-596.

Lebenthal, E. (1975) Cows milk protein allergy. Pediatr. Clin. North Am. **22** 827-833.

Liberatori, J., Morisio, L.,Guidetti, I.D. & Conti, A. (1979) Immunological evidence of ß-lactoglobulins in human colostrum and milk. Boll. Soc. It. Biol. Sper. **55** 822-825.

Liberatori, J., Morisio, L.,Guidetti, I.D, Conti, A. & Napolitano, L. (1979) ß-lactoglobulins in the mammary secretions of camel (*Camelus dromedarius*) and she-ass. Immunological detection and prelimenary physico-chemical characterisation Boll. Soc. It. Biol. Sper. **55** 1369-1373.

Lopez, C., Grubb, A., Soriano, F. & Mendez, E. (1981) The complete amino acid sequence of human complex forming glycoprotein heterogenous in charge (Protein HC). Biochim. Biophys. Res. Com. **98** 997-1001.

Lyster, R. (1972) reviews of the progress of Dairy science. Section C. Chemistry of milk proteins. J. Dairy Res. **39** 279-318.

McAlpine, A.S. & Sawyer, L. (1990) β–lactoglobulin: a protein drug carrier? Biochem. Soc. Trans. **18**, 879.

McKenzie, H.A. (1971) β–lactoglobulins. Milk Proteins **II** 257-330.

McKenzie, H.A., Ralston, G.B. & Shaw, D.C. (1972) Location of sulphydryl and

disulphide groups in bovine ß-lactoglobulins and effects of urea. Biochemistry **11** 4539-4547.

McKenzie, H.A. & Sawyer, W.H. (1967) Effect of pH on β–lactoglobulins. Nature **214** 101-1104.

McMeekin, T.L., Polis, B.D., Dellamonica, E.S. & Custer, J.H. (1949) A crystalline compound of β–lactoglobulin with dodecyl sulphate. J. Am. Chem. Soc. **71** 3606-3609.

McPherson, A. (1976) Preparation and analysis of protein crystals. Wiley Interscience.

McPherson, A. (1976) Crystallisation of proteins from polyethylene glycol. J. Biol. Chem. **251** 6300-6303.

McPherson, A. (1989) Macromolecular crystals. Scientific American **260** (March) 42-49.

Malik, Z., Bottomly, R. & Austen, B. (1988) Allergic properties of the genetic varients A and B of bovine ß-lactoglobulin. Int. Arch. All. Appl. Immun. **86** 245-248.

Mallinson, P.D. & Muir, K.W. (1985) ORTEP II, Interactive version. J. Appl. Cryst. **18** 51-53.

Marcon-Genty, D., Tome, D., Kheroua, O., Dumontier, A-M., Heyman, M. & Desjeux, J-F. (1989) Transport of β–lactoglobulin across rabbit ileum in vitro. Am. J. Physiol. **256** G943-G948.

Michel, H. & Oesterhelt, D. (1980) Three dimensional crystals of membrane proteins: bacteriorhodopsin. Proc. Natl. Acad. Sci. **77** 1283-1285.

Miranda, G. & Pelissier, J.P. (1983) Kinetic studies of in-vivo digestion of bovine unheated skim-milk proteins in the rat stomach. J. Dairy Res. **50** 27-36.

Monaco, H.L., Zanotti, G., Spadon, P., Bolognesi, M., Sawyer, L. & Eliopoulos, E.E. (1987) Crystal structure of the trigonal form of β–lactoglobulin and its complex with retinol at 2.5 A resolution. J. Mol. Biol. **197** 695-706.

Monti, J.C., Mermoud, A.C. & Jolles, P. (1989) Anti-bovine β–lactoglobulin antibodies react with a human lactoferrin fragment and bovine β–lactoglobulin present in human milk. Experentia **45** 178-180.

Motherwell, W.D.S. (1972) PLUTO. Program for plotting molecular and crystal structures. University of Cambridge, England.

Newcomer, M.G., Jones, T.A., Aqvist, J., Sundelin, J., Erikkson, U., Rask, L. 7 Peterson, A. (1984) The three dimensional structure of retinol-binding protein.

EMBO J. **3** 1451-1454.

North, A.C.T. (1989) Three-dimensional arangement of conserved amino acid residues in a superfamily of specific ligand binding proteins. Int. J. Biol. Macromol. **11** 56-58.

Otani, H., Uchio, T. & Tokita, F. (1985) Antigenic reactivities of chemically modified ß-lactoglobulins with antiserum to bovine ß-lactoglobulin. Agri. Biol. Chem. **49** 2531-2536.

Palmer, A.H. (1934) The preparation of a crystalline globulin from the albumin fraction of cow's milk. J. Biol. Chem. **104** 359-372.

Papiz, M.Z. (1982) Crystallographic studies of β–lactoglobulin at 2.8 A resolution. Ph.D. thesis, Napier college, Edinburgh.

Papiz, M.Z., Sawyer, L., Eliopoulos, E.E., North, A.C.T., Findlay, J.B.C., Sivaprasadarao, R., Jones, T.A., Newcomer, M.A. & Kraulis, P.J. (1986) The structure of β–lactoglobulin and its similarity to Plasma retinol binding protein. Nature **324** 383-385.

Paterson, G. (1991) The expression and secretion of ovine ß-lactoglobulin in yeast. Ph.D. thesis, University of Edinburgh.

Pierce, A.E. (1960) β–lactoglobulins in the urine of the new-born suckled calf. Nature **88** 940-941.

Perez, M.D., Diaz de Villegas, C., Sanchez, L., Arando, P., Ena, J.M. & Calvo, M. (1989) Interaction of fatty acids with ß-lactoglobulin and albumin from ruminant milk. J. Biochem. **106** 1094-1097.

Peroutka, S.J. (1988) 5-HT1A Receptors. Ann. Rev. Neurosci. **11** 45-60.

Pervaiz, S. & Brew, K. (1985) Homology of ß-lactoglobulin, serum retinol-binding protein and protein HC. Science **228** 335-337.

Plaxton, W.C. & Moorhead, G.B.G. (1989) Peptide mapping by CNBr fragmentation using a sodium dodecyl sulphate polyacrylamide minigel system. Anal. Biochem. **178** 391-393.

Rossmann, M.G. (1972) The molecular replacement method. (editor) Int. Sci. Rev., Gordon & Breach.

Rossmann, M.G. (1990) The molecular replacement method. Acta Cryst **A46** 73-82.

Said, H.M., Ong, D.E. & Shingleton, J.L. (1989) Intestinal uptake of retinol: enhancement by bovine milk β–lactoglobulin. Am. J. Clin. Nutr. **49** 690-694.

Savilhati, E. (1981) Cows milk allergy. Allergy **36** 73-76.

Sawyer, L. & Richardson, J. (1991) letter in TIBS **16** 11.

Sawyer, L., Papiz, M.Z.; North, A.C.T. & Eliopoulos, E.E. (1985) Structure and function of bovine β–lactoglobulin. Biochem. Soc. Trans. **13** 265-266.

Schagger, H. & von Jagow, G. (1987) Tricine - sodium dodecyl sulphate - polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem. **166** 368-379.

Seibles, T.S. (1969) Interaction of dodecyl sulphate with native and modified β–lactoglobulin. Biochemistry **8** 2949-2953.

Shaw, D. (1991) personal communication.

Sheldrick, G. (1986) SHELXS86. Program for crystal structure solution. University of Gottingen, Federal Republic of Germany.

Sheriff, S., Silverton, E.W., Padlan, E.A., Cohen, G.H., Smith-Gill, S.J., Finzel, B.C. & Davies, D.R. (1987) Three-dimensional structure of an antibody-antigen complex. Proc. Natl. Acad. Sci. **84** 8075-8079.

Siurala, M., Mustala, O. & Jussila, J. (1969) Absorption of acetylsalacylic acid by an normal and an atrophic gastric mucosa. Scand. J. Gastroent. **4** 269-273.

Steinrauf, L.K. (1959) Preliminary X-ray data for some new crystalline forms of β–lactoglobulin and hen egg-white lysozyme. Acta Cryst. **12** 77-78.

Stewart, D.E., Sarker, A. & Wampler, J.E. (1990) Occrence and role of *cis* peptide bonds in protein structures. J. Mol. Biol. **214** 253-260.

Tanford, C., Bunville, L.G. & Nozaki, Y. (1959) The reversible trans-formation of ß-lactoglobulin at pH 7.5. J. Am. Chem. Soc. **81** 4032-4036.

Tanford, C. & Taggart, V.G. (1961) Ionization-linked changes in protein conformation. II. the N - R transition in ß-lactoglobulin. J. Am. Chem Soc. **83** 1634-1638.

Tilley, J.M.A. (1960) The chemical and physiological properties of bovine ß-lactoglobulin. Dairy Sci. Abs. **22** 111-125.

Timasheff, S.N. & Townend, R (1961) Molecular interactions in ß-lactoglobulin. IV. The association of the genetic species of ß-lactoglobulin below the isoelectric point. J. Am. Chem. Soc. **83** 464-469.

Timasheff, S.N., Townend, R. & Mescanti, L. (1966) The optical rotatory dispersion of β–lactoglobulins. J. Biol. Chem. **241** 1863-1870.

Totsuka, M., Katakura, Y., Shimizu, M., Kumagai, I. Miura,, K-i & Kaminogawa, S. (1990) Expression and secretion of bovine ß-lactoglobulin in *Saccharomyces cerevisiae*. Agric. Biol. Chem. **54** 3111-3116.

Townend, R., Herskivits, T.T., Timasheff, S.N. & Gorbunoff, M.J. (1969) The state of amino acid residues in β–lactoglobulin. Arch. Biochem. Biophys. **129** 567-580.

Townend, R., Kumosinski, T.F. & Timasheff, S.N. (1967) The circular dichroism of variants of β–lactoglobulin. J. Biol. Chem. **242** 4538-4545.

Townend, R. & Timasheff, S.N. (1960) Molecular interactions in ß-lactoglobulin. III. Light scattering investigation of the stoichiometry of the association between pH 3.7 and 5.2. J. Am. Chem. Soc. **82** 3168-3174.

Townend, R., Weinberger, L. & Timasheff, S.N. (1960) Molecular interactions in β–lactoglobulin. IV. The dissociation of β–lactoglobulin below pH 3.5. J. Am. Chem. Soc. **82** 3175-3179.

Tronrud, D.E., Monzingo, A.F. & Mathews, B.W. (1986) Crystallographic structural analysis of phophoramidates as inhibitors and transition state analogs of thermolysin. Eur. J. Biochem. **157** 261-268.

Tronrud, D.E., Ten Eyck, L.F. & Mathews, B.W. (1987) An efficient general-purpose least squares refinement program for macromolecular structures. Acta Cryst. **A43** 489-501.

Udall, J.N., Colony, P., Fritze, L., Pang, K., Frier, J-S. & Walker W.A. (1981) Development of gastrointestinal mucosal barrier. II. The effect of natural versus artificial feeding on intestinal permeability to macromolecules. Pediatr. Res. **15** 245-249.

Udall, J.N., Pang, K., Fritze, L., Klienman, R. & Walker W.A. (1981) Development of gastrointestinal mucosal barrier. I. The effect of age on intestinal permeability to macromolecules. Pediatr. Res. **15** 241-244.

Verlet, L. (1967) Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard Jones molecules. Phys. Rev. **159** 98-103.

Weis, W.I. & Brunger, A.T. (1989) Crystallographic refinement by simulated annealing. Proc. Daresbury study weekend 16-28.

Wilson, N.H. & Jones, R.L. (1985) Prosatglandin endoperoxide and thromboxane A2 analogs. Thromboxane Leukotrine Res. **14** 393-425.

Yewdall, S.J. (1988) Structural studies of β–lactoglobulin. Ph.D. thesis, University of Leeds.

Zeppezauer, M., Soderberg, B.O. & Branden, C.I. (1967) Crystallisation of horse liver alcohol dehydrogenase complexes from alcohol solutions. Acta Chem. Scand. **21** 1099-1101.

APPENDIX 1

Appendix 1.

A1.1 Introduction.

A correlation between the species distribution of Blg and placental type of that species was proposed by Holt & Sawyer (personal communication) and by Yewdall (1988). The distribution of Blg covers a variety of species and it was thought that this might lead to a biological function being attributed to the protein. The distribution of Blg is covered in Appendix 2. Blg has now also been shown to exist in cat (Halliday *et al.*, 1991) and also in macaque (Shaw, personal communication) both of which are mammals

Blg is a milk specific protein and, therefore, its presence is likely to be important for a suckling animals health soon after parturition. The apparent absence of Blg, if it is assumed to have a specific function other than nutrition, in some species must be compensated for by other means.

On examination of the presence of Blg and the complexity of interaction between the foetal and maternal tissues during pregnancy a relationship seemed to be apparent. The more intimate the relationship between the two circulations the less likely that Blg was present in the maternal milk.

Placental types are classified in five basic groups according to the degree of intamacy of the chorion with the uterine mucosa (see Table A1.1). The lemur and chimpanzee fall into the group which has the least intimate placental union, epithelio-chorial, a category which is correlated with the secretion of Blg within the milk. With the lemur and chimpanzee being primates, an analogy with man would however suggest that they would not be expected to secrete Blg.

Placental type	Epithelio-chorial	Syndesmo-chorial	Endothelio chorial	Haemo-chorial	Haemo-Endothelial
MATERNAL TISUUES					
Endothelium	+	+	+	-	-
Connective tissue	+	+	-	-	-
Epithelium	+	-	-	-	-
FOETAL TISSUE					
Chorionic epithelium	+	+	+	+	-
Mesenchyme	+	+	+	+	-
Endothelium	+	+	+	+	+
EXAMPLES	Horse Pig	Cattle Sheep	Dog Cat	Insectivores Lower rodents Man	Rat Rabbit
Blg SECRETION	+	+	+	-	-

Table A1.1 Summary of placental types.

+ indicates presence, - indicates absence

To investigate whether these two animals had Blg present within their milk, their whey fractions were separated according to Davies (1974) at the Hannah Institute, Ayr (Figure A1.1). These fractions were then examined here by gel electrophoresis and by immunoblotting using antibovine Blg antisera.

A1.2 Materials and Methods.

Please refer to Chapter 4 for the materials and methods for Immunoblots and polyacrylamide gel procedures. These are identical to those carried out in this section.

A1.2.1 Succinylation of Fraction 3 from Lemur Milk.

1mg of fraction 3 from lemur milk whey separation was weighed into 0.2 ml of 2M TRIS. The solution was mixed and a small amount of succinic anhydride added on the end of a microspatula and dissolved. This procedure was repeated every 30 minutes up to 4 hours, after which the mixture was dialysed against several changes of 1% ammonium bicarbonate for 4 hours.

A1.2.2 Clostripain Digestion.

The enzyme clostripain (2mg/ml) was activated by preincubation in 1mM calcium acetate and 2mM dithiothreitol (DTT) overnight at 4 C. To 100ul of dialysed succinylated sample, 10ml of 75mM DTT solution in 1% ammonium carbonate was added and mixed. To this solution 5ml of activated clostripain was added and placed in a 37 °C waterbath. After 4 hours incubation the the sample was dried in a speedvac vacuum centrifuge.

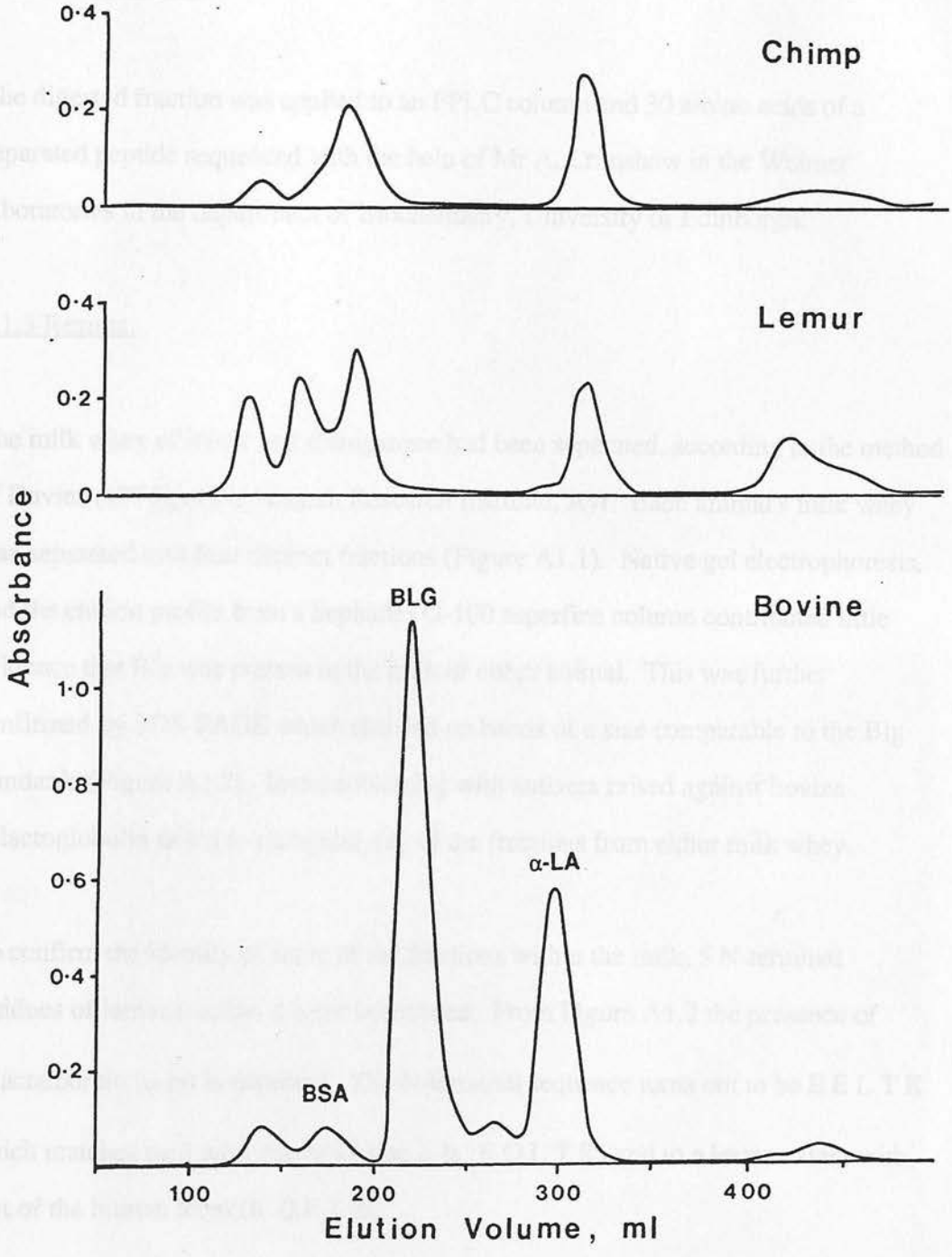


Figure A1.1 Chromatograms from a G100 superfine fractionation for bovine, chimp and lemur milk whey.

A1.2.3 Sequencing.

The digested fraction was applied to an FPLC column and 30 amino acids of a separated peptide sequenced with the help of Mr A. Cronshaw in the Welmet laboratories in the department of Biochemistry, University of Edinburgh.

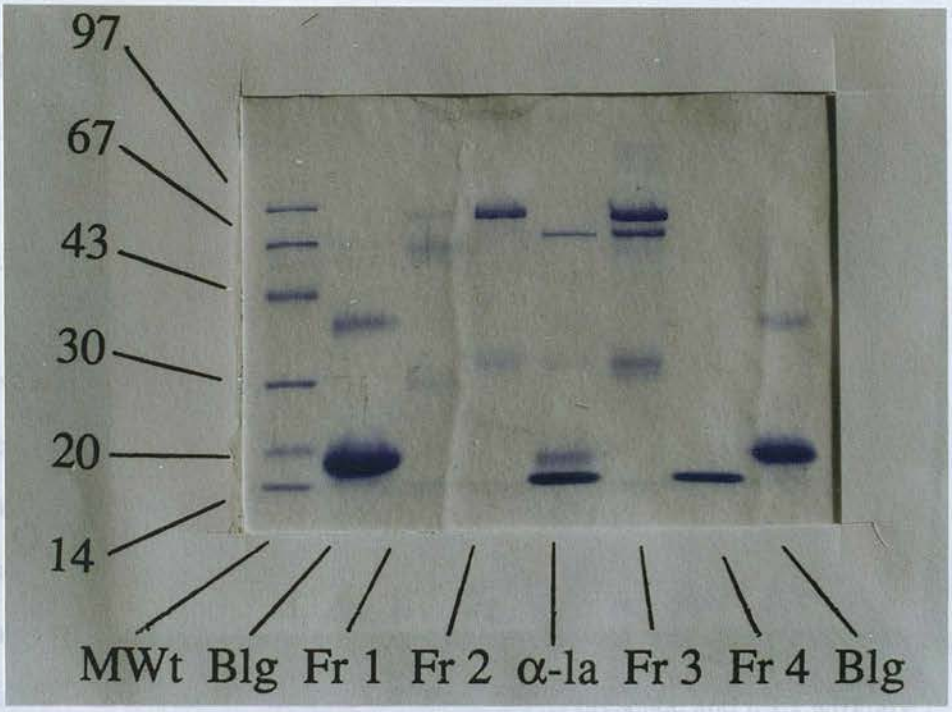
A1.3 Results.

The milk whey of lemur and chimpanzee had been separated, according to the method of Davies (1974), at the Hannah Research Institute, Ayr. Each animal's milk whey was separated into four distinct fractions (Figure A1.1). Native gel electrophoresis, and the elution profile from a Sephadex G-100 superfine column contributed little evidence that Blg was present in the milk of either animal. This was further confirmed by SDS-PAGE which showed no bands of a size comparable to the Blg standards (Figure A1.2). Immunoblotting with antisera raised against bovine B-lactoglobulin failed to recognise any of the fractions from either milk whey.

To confirm the identity of some of the fractions within the milk, 5 N-terminal residues of lemur fraction 4 were sequenced. From Figure A1.2 the presence of α-lactalbumin (α-la) is expected. The N-terminal sequence turns out to be E E L T K which matches well with that of bovine α-la (E Q L T K) and to a lesser extent with that of the human form (K Q F T K).

N-terminal sequencing of Lemur fraction 3 was hindered by it being blocked. As a result the protein was succinylated before clostripain digestion, to limit sites of cleavage to arginine residues. After digestion, peptides were separated and the first 30 amino acids of one peptide sequenced. The sequenced peptide shows 76.7%

(a)



(b)

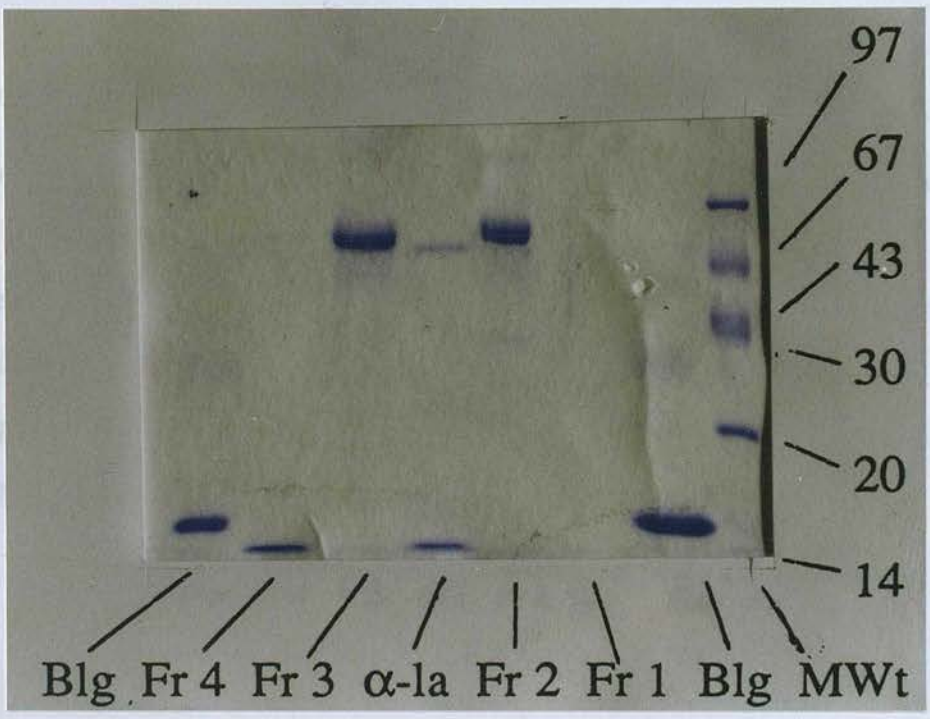


Figure A1.2 Gel electrophoresis patterns of fractions 1 to 4 from chimp (a) and lemur (b) milk whey fractionation

MWt standards are x 1000 D.

identity with residues 26 - 56 of bovine preproalbumin with a further 7 out of the 30 residues having conservative changes (Figure A1.3).

A1.4 Conclusions.

The evidence presented here suggests that there is no Blg in the milk of either chimpanzee or lemur. If a Blg homologue is indeed present in the milk of all species then its non detection may be a result of it being present at such low levels. This seems unlikely since antibody tests are very sensitive and should pick up the protein even at low levels. It may also be possible that the protein in some species becomes associated with another in the milk which masks its presence and tests wrongly indicate that the protein is in fact absent.

The tests here also show the closer relationship of the two proteins sequenced, namely α-la and serum albumin, with their bovine counterparts rather than the human forms. The lemur and chimpanzee are however similar to the rest of the primates in that they apprently contain no Blg. It disproves the theory that the the intimacy of contact between mother and offspring leads to the prediction of presence of Blg. Although placental anatomy may provide a good guide to whether the secretion of Blg (or a homologue) is expected, it is not a fool proof method.

```
seq. 1                                    D T H K S E L A H R Y N D L G E E H F
                                          | | | | | | | : | | | : | | | | | |
seq. 2    M K W V T F I S L L L L F S S A Y S R G V F R R D T H K S E I A H R F K D L G E E H F

seq. 1    K A L V L V T F S Q F
          | : | | | | : : | | | :
seq. 2    K G L V L I A F S Q Y L Q Q C P F D E H V K L V N E L T E F A K T C V A D E S H A G C E

seq. 2    K S L H T F G D E L C K V A S L R A T Y G D M A D C C E K E P
```

Figure A1.3 Sequence allignment of Lemur Fr. 3 and pre-pro albumin.

seq.1 = 30 residue peptide of lemur fraction

seq.2 = bovine pre-pro albumin sequence.

APPENDIX 2

ß-Lactoglobulin

Susan G Hambling, Alan S McAlpine

and Lindsay Sawyer*

Department of Biochemistry,
University of Edinburgh Medical School,
Hugh Robson Building,
George Square,
Edinburgh EH8 9XD,
Scotland

* To whom proofs and correspondence should be addressed

Introduction

ß-Lactoglobulin (BLG) is the major whey protein secreted in the milk of ruminants like the cow or sheep, and monogastrics like the pig, horse, dog or cat. It is absent from the milk of humans and rodents. Bovine BLG can be readily isolated and, since its isolation from milk by Palmer in 1934, it has been extensively studied by essentially every technique available. Despite this, however, its biological function still remains uncertain.

The name ß-lactoglobulin was first proposed by Cannan *et al.* (1942) for the fraction crystallised by Palmer, which Pedersen (1936a) had identified with the 'ß' component of whey in ultracentrifugation studies. Despite its having been called 'ß-lactalbumin' by Svedberg (1938), the name ß-lactoglobulin is that by which the protein is now known and it is *quite distinct* from the other whey protein α-lactalbumin, the subject of another chapter in this volume!

A number of reviews on the physico-chemical properties of BLG have been published over the years (*eg.* Tilley, 1960; Townend *et al.*, 1969; McKenzie, 1971; Green *et al.*, 1979) together with several others on milk proteins in general which contain substantial sections on the properties of the protein (*eg.* McKenzie, 1967; Lyster, 1972; Thompson & Farrell, 1974; Whitney *et al.*, 1976; Jenness, 1979; Swaisgood, 1982; Jenness, 1985). It is the purpose of this present review to bring together various strands of the work on ß-lactoglobulin over the last 55 years and to discuss the various clues that these give to the function of BLG.

For example, binding studies carried out on the bovine protein *in vitro* have shown that it can bind a variety of ligands, most of which are small, hydrophobic molecules like fatty acids or vitamin A (Futterman & Heller, 1972). ORD studies on BLG have shown that it undergoes several conformational changes between pH 2 and pH 9, possibly the most important of which is the N <-> R transition occurring in the physiological pH range (Tanford *et al.*, 1959). Sedimentation studies reveal the bovine protein to be a dimer under physiological conditions whereas the BLG from pigs is a monomer. Crystal structures of BLG at both pH 6.5 and pH 7.8 have recently become available and are being improved so that it should soon be possible to offer a detailed molecular explanation for these and other solution properties.

The primary structure of BLG shows some homology to human serum retinol binding protein (RBP) (Godovac-Zimmerman *et al.*,1985a; Pervaiz & Brew, 1985). More striking, however, is the similarity in tertiary structure (Sawyer *et al.*, 1985; Papiz *et al.*, 1986) which also extends to insecticyanin (ICN) (Holden *et al.*,1987; Sawyer, 1987). Both of these, RBP and ICN, carry small insoluble molecules (the former retinol, the latter biliverdin) within a hydrophobic cavity, which is characteristic of all three of these proteins. Further similarities, at least at the sequence level, have been identified in a number of other proteins giving rise to the likelihood that there is a family of proteins, the lipocalycins (*cf.* lipocalins, Pervaiz & Brew, 1987), which transport small, insoluble or labile molecules. That they may all have evolved from a common ancestor is supported by recent work on the gene sequences which show a similar intron/exon pattern for the several members of the family for which these data are available.

In what follows, it is not our intention to include any of the extensive literature on the food science aspects of the protein since this will be covered elsewhere. Rather, we shall concentrate

upon the studies which have been carried out under what are basically physiological conditions and which provide information about BLG as an intriguing biological macromolecule.

Biosynthesis and Secretion

Major lactoproteins, including BLG, are biosynthesised within the secretory epithelial cells of the mammary gland under the control of prolactin (Larson, 1979). mRNA coding for BLG which is specific to the mammary tissue, is translated to yield a 180 amino acid pre-BLG. Pre-BLGs have been isolated from the mammary glands of cow (Yoshikawa *et al.*, 1978), sheep and pig, using cell-free translation systems (Mercier & Gaye, 1982). The signal peptides consist of 16 or 18 highly conserved amino acids, which are hydrophobic, and possibly α- helical, with a basic residue near the N-terminus.

Cow Met-Lys-Cys-Leu-Leu-Leu-Ala-Leu-Ala-Leu- -Thr-Cys-Gly-Ala-Gln-Ala-
Sheep Met-Lys-Cys-Leu-Leu-Leu-Ala-Leu-Gly-Leu-Ala-Leu-Ala-Cys-Gly-Val-Gln-Ala-
Pig Met-Arg-Cys-Leu-Leu-Leu-Thr-Leu-Gly-Leu-Ala-Leu-Leu-Cys-Gly-Val-Gln-Ala-

A membrane-bound proteinase cleaves off the signal peptide on the C-terminal side of Ala, to generate the mature 162 amino acid protein which can then undergo post-translational modifications which begin in the cisternal space of the rough endoplasmic reticulum. Transfer of an oligosaccharide from membrane lipid to Asn in the protein sequence -Asn-X-Ser/Thr- has only been reported for BLG from the Australian Droughtmaster breed of cattle (Bell *et al.*,1970). It is likely that a disulphide isomerase ensures correct folding (Freedman *et al.*, 1989) before transport to the Golgi apparatus and subsequent incorporation into secretory vesicles. The native protein is then secreted into the lumen, where it accumulates in the milk before removal by the suckling young.

Distribution

The composition of milk varies considerably from species to species and also with the time since parturition (Larson, 1979). There are also seasonal variations, which presumably reflect the dietary habits of the animal. A quantitative partition of the albumin fraction of whey proteins useful for detecting the presence of BLG, has been devised by Davies (1974).

Since the initial preparation of BLG from bovine (*Bos taurus*) milk (Palmer,1934) dimeric BLGs have been isolated from the milks of a number of other ruminants, and monomeric BLGs have been purified from the milks of several non-ruminant livestock species (**TABLE I**). BLG has also been detected in milks of other species, but the state of its association in these cases is uncertain. For instance, the molecular weight of kangaroo BLG has been quoted as 17 to 19kDa (McKenzie *et al.*, 1983), suggesting a monomeric structure but a subsequent report declared that it was dimeric, despite its low (32%) sequence homology with bovine BLG (Godovac-Zimmerman & Shaw, 1987).

It seems likely that any BLGs detected by bovine BLG antiserum are dimeric, as there is known to be little cross-reactivity between this antiserum and monomeric BLGs. Hence the lack of cross-reactivity does not rule out the presence of a monomeric BLG. Thus, in two cases, donkey (Godovac-Zimmerman *et al.*, 1988) and peccary (Lyster *et al.*, 1966; Kalan *et al.*, 1971) there was no cross-reactivity with bovine BLG antiserum whereas there was when antiserum to monomeric BLG was used. The existence of camel BLG is still in doubt: its absence was stated (quoted by Kessler & Brew, 1970), and it is known that camel milk does not cross-react with antiserum to bovine BLG (Lyster *et al.*, 1966). However, higher titres of antibodies to bovine BLG have subsequently been shown to cross-react with both camel milk and colostrum (Liberatori *et al.*, 1979b). This may indicate that there is a monomeric BLG present. Preliminary studies with feline milk revealed there to be no cross-reactivity between it and bovine BLG antiserum, but recently a protein thought to be BLG has been isolated and is being sequenced (H.A.McKenzie, personal communication, 1989). It is probable, that this BLG will turn out to be monomeric.

There is believed to be no BLG in human milk (Bell & McKenzie, 1964) despite the isolation of a 13.5 kDa whey protein with a pI of 5.1 (Conti *et al.*, 1980), and the characterisation of a protein from human colostrum with a pI of 6.1, a molecular weight around 18kDa and an amino acid composition similar to that of pig BLG (Conti *et al.*, 1982b). Both of these proteins were detected by immunoprecipitation with bovine BLG antiserum, and were most likely fragments of human lactoferrin which is known to cross-react with the antiserum (Brignon *et al.*, 1985). Radioimmunoassays have suggested that low concentrations of bovine BLG (5-800 μg/l), or at least antigenic peptides there from, are present in human milk having been absorbed from the diet (Axelsson *et al.*, 1986; Monti *et al.*, 1989).

BLG is also absent from the milk of rodents. However, both rat and mouse milks contain a major whey acidic protein (WAP) which has a similar molecular weight to, but is not, a BLG. There is no evidence of hybridization between cDNA from a rat mammary library and cDNA for sheep BLG (Ali, 1989), and mouse WAP shows no primary sequence homology to bovine BLG (Hennighausen & Sippel, 1982). Despite a report that the major whey protein in murine milk (Lockwood *et al*, 1966) was BLG, it failed to cross-react with bovine BLG antisera (Lyster *et al*, 1966) and was most likely WAP. A recent publication has confirmed both the abundance of WAP (2g/l) and the absence of BLG (Clark *et al*, 1987). The 'mouse BLG' that has become commercially available is probably WAP, despite the successful expression of sheep BLG in transgenic mice (Simons *et al.*, 1987).

Isolation

The isolation of BLG from milk is a simple procedure, involving just four stages: the removal of fats, the removal of the caseins, the fractionation of the whey proteins, and the purification of BLG. Since each stage can be carried out in a number of ways, a variety of published methods are available.

The original method of Palmer (1934) for the isolation of bovine BLG from skim milk was eventually superceded by the method of Aschaffenburg & Drewry (1957), which involved an acid precipitation step at pH 2. Mailliart and Ribadeau-Dumas (1988) have devised a similar salting out procedure suitable for large scale use. Later, it was suggested that the harshness of this step might cause irreversible changes to BLG and the method was modified by Armstrong *et al.* (1967) such that the whey proteins were fractionated by acid precipitation at pH 3.5. More recently, Monaco *et al.* (1987) separated the whey proteins by DEAE-cellulose chromatography without the pH going below 6.6. Crystals grown from this BLG diffracted to a higher resolution than those grown from BLG purified by the method of Aschaffenburg & Drewry, suggesting that acid precipitation at pH 2 does indeed damage the protein in some undefined way. Despite this, most commercial preparations of bovine BLG use the procedure of Aschaffenburg & Drewry. A summary of the these methods is presented in **TABLE II.**

Other dimeric BLGs have also been obtained by these methods, or slight variations thereof. Yak BLG (Grosclaude *et al.*, 1976) has been isolated by the procedure of Aschaffenburg & Drewry, whereas red deer (McDougall & Stewart, 1976), waterbuffalo (Braunitzer *et al.*, 1979), sheep (Maubois *et al.*, 1965) and goat BLGs (Kalan & Basch, 1969) have been obtained by a modified method. More recently both goat (Preaux *et al.*, 1979) and ovine (Godovac-Zimmerman *et al.*, 1987) BLGs have been isolated by a non-salt precipitation method in which the fat and caseins were removed by centrifugation, and then the whey protein mixture was dialysed, lyophilized, and separated by gel filtration (cf. Davies, 1974). A procedure for the isolation of BLG from both cow and goat milk based upon ethanol fractionation has been reported (Bain & Deutsch, 1948) but neither isolate crystallised suggesting that some form of denaturation had occurred.

The isolation of monomeric BLGs from the milk of non-ruminants can be carried out with equal ease using procedures similar to those described above. Pig BLG has been isolated and purified by three different methods, which are summarized in **TABLE III**. No comparison appears to have been published on the advantages of these various methods. The procedure of Kalan *et al.* (1971) uses many precipitation steps whiche probably cause of the low yield (25%) obtained. The method of Kessler and Brew (1970), on the other hand, has been successfully used to isolate monomeric BLGs from the milks of dolphin, manatee, beagle (Pervaiz & Brew, 1986) and horse (Godovac-Zimmermann *et al.*, 1985b).

The amounts of BLG obtained by the various methods depend upon both the procedures used, and the quantity of BLG in the initial milk, which is known to vary with season and time since parturition as well as upon the isolation procedure used. The quantities of BLG isolated from the milks of a few species are given in **TABLE IV.**

Genetic Variants and Primary Structure

Although Li (1946) demonstrated by electrophoresis that BLG was heterogeneous, and Polis *et al.* (1950) separated two components of BLG by fractional crystallisation, it was Aschaffenburg & Drewry (1955) who showed that there were two components in bovine milk whose origin was probably genetically determined, a fact that has been proved (Aschaffenburg & Drewry; 1957) and reviewed (Aschaffenburg, 1968; see also Chapter **). The presence of genetic variants of BLG has also been shown in other species. The original phenotyping procedure used paper electrophoresis at pH 8.6 but a more convenient method uses gel electrophoreis on undiluted, whole milk samples (Aschaffenburg, 1964; Aschaffenburg & Thymann, 1965; Thompson, 1970; Kiddy *et al.*, 1972).

The original fractionation of the whey proteins of bovine milk produced a heterogeneous preparation of BLG. This was due to the presence of two genetic variants, BLGs A and B, which have been separated (Piez *et al.*, 1961), sequenced (Braunitzer *et al.*, 1972), and are the ones most often present in commercial preparations. Other genetic variants of bovine BLG have also been detected, isolated, and sequenced. Cow (*Bos taurus*) is currently believed to have six variants: BLG-A and BLG-B as mentioned above, BLG-C which has only been positively identified in the milk of Jersey cows (Bell & McKenzie, 1967a), but may also be present in South African Nguni cattle (Osterhoff & Pretorius, 1966), BLG-D (Brignon *et al.*,1969; Brignon & Ribadeau-Dumas, 1973), BLG-H (Conti *et al.*, 1988) and BLG-W which is thought to exhibit only neutral amino acid substitutions when compared to BLG-B (Godovac-Zimmermann *et al.*, 1990). A cross-breed between *Bos taurus* and *Bos indicus*, Australian Droughtmaster, was shown to produce variants BLG-A, BLG-B and a new variant BLG-Dr which is the only BLG known to be N-glycosylated (Bell *et al.*, 1970). Bali cattle, *Bos javanicus*, have three genetic variants. It has recently been suggested that BLG-E is identical to yak (*Bos grunniens*) BLG, as position 11 is now thought to be Asp (Bell *et al.*, 1981b) and not Asn (Grosclaude *et al.*, 1976). These data imply, not unexpectedly, a close relationship between bovine subgenera.

Other ruminant BLGs also exhibit genetic polymorphism. Red deer possess two variants (McDougall & Stewart, 1976), whereas sheep have variants A, B (Bell & McKenzie, 1967b; Preaux *et al.*, 1980; Kolde & Braunitzer, 1983) and C (Erhardt *et al.*, 1989). Mouflon BLG is homogeneous, and identical in sequence to ovine BLG-B (Godovac-Zimmerman *et al.*, 1987). Waterbuffalo (Braunitzer *et al.*, 1979) and goat (Sen & Chaudhuri, 1962; Preaux *et al.*, 1979) BLGs are both homogeneous by electrophoresis, but the latter (unlike any of the above proteins), contains an amino acid substitution which cannot be the result of a single point mutation, namely Asp-130 (GAU/C) in bovine BLG-A becomes Lys-130 (AAA/G) in goat.

Monomeric BLGs from kangaroo, dolphin, horse and pig, have also been purified to homogeneity and sequenced. Their sequences diverge considerably from that of bovine BLG-A and from one another. Many of the substitutions observed cannot have arisen from single point mutations. Kangaroo BLGs are homogeneous (Godovac-Zimmerman & Shaw, 1987), whereas dolphin, donkey, horse and pig show genetic polymorphism. Two dolphin BLGs have been detected and partially sequenced (Pervaiz & Brew, 1985,1986). Of the two donkey BLGs isolated, one has been fully sequenced (Godovac-Zimmerman *et al.*, 1988). Equine BLG has been shown to be heterogeneous (Conti *et al.*, 1982a), despite an earlier report to the contrary (Bell *et al.*, 1981a). Two variants have been purified and sequenced - the complete primary structure of BLG-I (Conti *et al.*, 1984) agrees with its partial sequence determined earlier (Bell *et al.*, 1981a), but reveals only 70% identity to BLG-II (Godovac-Zimmerman *et al.*, 1985a)

.

The heterogeneity of pig BLG was first reported in 1969 (Kraeling & Gerrits, 1969) and the variants named BLG-A and BLG-B (Kalan *et al.*, 1971). A later study also isolated two variants: one was believed to be BLG-A, and the other, distinct from BLG-B, was named BLG-C (Bell *et al.*, 1981c). Recently BLG has been purified from sow's milk, and studied by two methods (Conti *et al.*, 1986b): chromatofocussing yielded three peaks of which only the first at pH 4.7, was homogeneous. This was sequenced as pig BLG-I (Conti *et al.*, 1986a). On the other hand, isoelectricfocusing with an immobilised pH gradient gave three bands, BLG-1 BLG-2 and BLG-3 with pIs of 4.4, 4.65 and 4.9 respectively. The amino acid compositions of all these proteins are given in **TABLE V** together with a composition from this laboratory from a pig believed to be BB. It is unclear which variants are the same. Homology between the first fifty amino acids of BLG-A (Bell *et al.*, 1981c) and BLG-I (Conti *et al.*, 1986a) suggests that these two may be equivalent. However, it is apparent from the published work on the pig BLGs that inadequate genetic phenotyping has been performed before the biochemical analysis.

Now that primary sequences are routinely obtainable, reviewing the literature on amino acid composition and the various N- and C-terminal analyses is considered unnecessary. The interested reader should refer to the full discussion by McKenzie (1971). One point worth reporting, however, is the removal of the two carboxy-terminal residues by carboxypeptidase A treatment of bovine variants A, B and C (Kalan *et al.*, 1965). The rate of removal of the residues from variant C is significantly slower than that from variants A or B. Of course the disulphide bridge in the native sequence terminates any further proteolysis so that a homogeneous product can be obtained. This has been crystallised (Greenberg & Kalan , 1965) and a low resolution crystal structure determined (Green *et al.*,1979).

Although an incomplete sequence was published by Frank & Braunitzer in 1967, the first full sequence was reported by Braunitzer *et al.* (1972). The positions of the disulphide bridges have given rise to some controversy. Whilst the cystine 66-160 has been unambiguously identified by a variety of methods (Mainferme *et al.* ,1971; Martial *et al.* ,1971; Perez-Gomez *et al.* ,1971 with revised numbering reported by Preaux & Lontie, 1972), McKenzie *et al.* (1972) have shown that equimolar amounts of 106-119 and 106-121 were obtained when ^{14}C-iodoacetamide was covalently bound to the free thiol in the presence of 8M urea, the disulphide bridges reduced, and then unlabelled iodoacetamide reacted with the newly generated sulphydryls. Subsequent peptide digestion and analysis showed that the label was evenly distributed between Cys-119 and Cys-121. However, McKenzie & Shaw (1972) report that there is no appreciable reaction between the free thiol and iodoacetamide in the absence of 8M urea, and that 8M urea can disrupt disulphide bridges (McKenzie *et al.*, 1972). Thus the interchange observed could be explained by the effects of urea. This result conflicts with those of Preaux & Lontie (1972), and the crystal structures at medium resolution (Papiz *et al.*, 1986; Monaco *et al.*, 1987; Yewdall, 1988) show clearly that there is but one disulphide linking residues 106-119. Indeed, the separation of the two Cys residues is some 10Å so that a significant rearrangement is required to accomplish the interchange. What is not ruled out by the crystallography is the selection by crystallisation of the 106-119 isomer, although the two forms which grow together from the same solution at pH 7.8, both have this cystine. Mutant proteins have been produced with either 119 or 121 replaced by Ser but they have not yet been analysed (Paterson, Fothergill-Gilmore & Sawyer, unpublished; Batt, personal communication, 1990).

Not only have there been a number of protein sequence determinations on a wide variety of BLGs (Godovac-Zimmermann & Braunitzer, 1987; Godovac-Zimmermann, 1988) but there are now a number of partial and complete DNA sequences available for mRNA-derived cDNA (Willis *et al.*, 1982; Mercier *et al.*, 1985; Gaye *et al.*, 1986; Jamieson *et al.*, 1987; Ivanov *et al.*,1988; Anderson *et al.*, 1989) and the gene (Ali & Clark, 1988) which reveals the presence of introns.

The genetic variants have usually been detected because they have a charge difference detectable on electrophoresis (but see Krause *et al.*, 1988). However, now that DNA sequencing techniques are available, mutations which are not reflected in altered mobility will perhaps be found more commonly. For example, the DNA sequence of bovine BLG determined by Anderson *et al.* (1989) implies the change Phe105Val (TTC->GTC). As no comment is made about this, however, an error of some kind cannot be ruled out.

Examination of the primary structure of BLG reveals no obviously repetitive or unusual stretches of sequence. The ruminant proteins as purified are not modified post-translationally, with the exception of the BLG-Dr whose Asp/Asn mutation has already been referred to. It is interesting to note however that they all possess a sequence Asn-Pro-Thr at 152 which does not appear to be glycosylated. From the crystal structure at pH 7.6 (Papiz *et al.*, 1986), this can be explained on steric grounds since the Asn residue lies in between the two subunits. Since assembly is assumed to occur as the last stage in the biosynthesis, this observation could mean that there is a penultimate *de*glycosylation process.

The BLG variants from ruminants show greater than 90% sequence identity between one another within the same species, and other dimeric BLGs (**TABLE VI**). They are synthesised under the control of codominant alleles. The non-ruminant variants are monomeric, show much lower primary sequence homology to other monomeric and dimeric BLGs (typically 30-70% identitiy), and may originate from gene duplication, as suggested for the horse BLGs (Godovac-Zimmerman *et al.*, 1985b). **TABLE VII** gives the sequences of the various non-ruminant species with cow BLG-A as the representative ruminant protein.

SOLUTION STUDIES

Probably because of its abundance, ease of preparation and relative stability, numerous solution studies have been carried out on this protein using a wide variety of techniques. These experiments have elucidated the physicochemical behaviour of the protein, provided an insight into the importance of specific amino acids and yielded information on the binding of ligands to BLG *in vitro*. The vast majority of these studies have been carried out with the bovine protein and, unless specifically mentioned, it is this protein which has been used. A number of physicochemical properties are summarised in **TABLE VIII**.

Solubility

Tilley reports the considerable variation in the solubility of BLG as reported in the literature up to 1960. In this laboratory, protein is taken up into solution in dilute buffer from which it can be precipitated and even crystallised by dialysis against distilled water (Senti & Warner, 1948). It is clear, however, that the addition of salt or buffer increases the solubility of BLG quite dramatically: Polis *et al.* (1950) dissolved 0.18g/100ml water at the pI compared to 1.65g/100ml at pH 5.2 in 0.2M NaCl, a tenfold increase. Similar increases were observed by Treece *et al.* (1964) who also showed BLG-B to be about 5 times more soluble than BLG-A. A solubility minimum always exists at the pI, and so different absolute values are obtained at other pHs. Arakawa & Timasheff (1987) have examined the behaviour of BLG with NaCl and conclude that because the solubility is anomolous, there is a considerable binding of electrolytes resulting from the unique distribution of surface charge at neutral pH, which results in the large observed dipole moment of 730 Debye (Ferry & Oncley, 1941). Salting-out from solutions containing well in excess of 50mg/ml protein is the standard way of growing crystals for X-ray studies (Aschaffenburg *et al.*, 1965)

Molecular Size

Ultracentrifugation studies by Pederson (1936b) found variation of the sedimentation coefficient with pH and, in sedimentation equilibrium studies found the M_w to increase from 37,600 at pH 6.4 to 50,200 at pH 9.8. This indicated a slow aggregation not observed during his sedimentation velocity measurements. The molecular weight determined by osmotic pressure (Gutfreund, 1945; Bull & Currie, 1946) , light scattering (Heller & Klevens, 1946; Halwer *et al.*, 1951), sedimentation and diffusion (Pederson, 1936b; Cecil & Ogston, 1949) and X-ray diffraction (McMeekin & Warner, 1942; Green *et al.*, 1956) gave values around 35,600. However, Bull found a value of 17,100 using a surface film technique, which further revealed a value of 34,300 when carried out in the presence of Cu^{++}. Subsequently, SDS-PAGE in reducing conditions and amino acid analysis provided values of about 18,500 (Piez *et al.*, 1961). Townend *et al* (1960) found a value of 18,000 at pH values below 3.5. These results are consistent with the normal form of ruminant BLG being a dimer of subunit molecular weight around 18,200, and this is entirely in keeping with the crystal structure work (Green & Aschaffenburg, 1959; Aschaffenburg *et al.*, 1965) and the value calculated from the sequence of 18,200. However, In the decade following 1957, a large number of studies were performed on the association of BLG as a function of pH, protein concentration and temperature. These studies are discussed fully

elsewhere (McKenzie, 1967,1971; Swaisgood, 1982). Here only the main conclusions are reported and updated as necessary.

As has been mentioned, below pH 3.5 dissociation into monomers occurs with little change to the ORD (Timasheff *et al.*, 1966a), CD (Townend *et al.*, 1967; Teller *et al.*, 1979) or tryptophan fluorescence (Townend *et al.*, 1967). Timasheff & Townend (1961) have examined the sedimentation and light scattering properties of the A and B variants and find dissociation constants at 15°C and pH 2.7 of 7.5×10^{-5} and 2.4×10^{-5} mol/l respectively. These values are not greatly affected by ionic strength (Townend *et al.*, 1960a,b,c) or genetic variant (Townend *et al.*,1964) and correspond to a free energy of association of about -25kJ/mol. In keeping with this, Mills & Creamer (1975) conclude that there is a rapid dissociation to monomer followed by a slow conformational change when the pH is dropped from 4.5 to 2.0, from both intrinsic and extrinsic fluorescence measurements .

Between pH 3.7 and 6.5 bovine BLG-A, and mixtures of A and B, can reversibly octamerize to produce aggregates of 144 kDa (Townend & Timasheff, 1960; Timasheff & Townend, 1961; Pessen *et al.*, 1985). Neither tetrameric nor hexameric intermediates are present in significant amounts. The interactions within the octamer of BLG-A are maximal around pH 4.5 and near 0°C. The association constant at pH 4.4, 4.5°C is $2.8 \times 10^{11} dm^3.mol^{-3}$, corresponding to a free energy of $-60.8 kJ.mol^{-1}$. The packing of subunits to form a compact octamer with 422 symmetry (Timasheff & Townend, 1964) is accompanied by an increase in hydration which proton NMR estimates to be 5-6 extra water molecules/dimer. Aromatic residues are believed to not be involved as their UV spectra are unaltered by association. However carboxyl groups have been implicated in this aggregation (Armstrong & McKenzie, 1967). Up to 30% modification of the carboxyls by carbodiimide has no effect on the ORD spectrum of the protein, but does reduce the protein's ability to octamerize. Two carboxyl residues likely to be involved are Asp-64, which is present in BLG-A but not BLG-B, the latter being unable to form an octamer alone, and Asp-28, which is the only amino acid substituted by Asn in BLG-Dr. Glycosylation at Asn-28 prevents octamerization of BLG-Dr, presumably by steric hindrance since the location of these residues in the structure shows them to be 24 Å apart.

Ruminant BLGs are dimeric over the pH range 5 to 8, although there is evidence that a monomer-dimer equilibrium exists in cow and red deer BLGs at pH 6.5 (McDougall & Stewart, 1977). The decrease in sedimentation coefficient observed across this pH range for bovine BLG-A and BLG-B most probably corresponds to an increase in the dissociation of the dimer (Georges *et al.*, 1962; McKenzie *et al.*, 1967). Around pH 5.3, which is near the isoelectric point for all bovine variants, the free energy of association is about -40kJ/mol (Timasheff & Townend, 1961; Reithel & Kelly, 1971). Zimmerman *et al.* (1970) found values for the bovine A and B variants at 20°C of -26.3 and -28.8 kJ/mol respectively for the free energies of association at pH 6.9, in fair agreement with the values of -23.6 and -28.6 kJ/mol found at pH 7.5 by McKenzie *et al.*, (1967) in ultracentrifugation studies and -23.6 kJ/mol by light scattering (Georges *et al.*, 1962) for variant B.

Over the pH range 8 to 9.5 slow time-dependent changes occur in BLG. At pH values above 8.5 reversible dissociation occurs (George & Guinand, 1960; Townend *et al.*, 1960b) and, at above pH 9 the optical rotation, circular dichroism and solubility alter with time as the protein denatures irreversibly and aggregates (Groves *et al.*, 1951; Christensen, 1952; Herskovits *et al.*, 1964). The addition of a thiol-blocking group can inhibit aggregation, implicating thiol oxidation and/or thiol/disulphide exchange in the formation of the heavier components (Roels *et al.*, 1966).

Preliminary crystallographic studies on the dimeric form of bovine BLG have suggested that the interface between subunits may involve both hydrophobic interactions between Ile-29 and Ile-147, and the stacking of the symmetry- related His-146 residues (Papiz *et al.*, 1986).

Little if any information about the aggregation of the monomeric BLG's has been published.

DENATURATION.

The denaturation of bovine BLG involves the dissociation of dimer to monomer, a major change in the conformation of the polypeptide chain, and aggregation. The denaturant can be alkali, heat, organic compounds or heavy metal ions, and it is not clear how each of these acts to yield insoluble aggregates. Although it is improbable that there is a common mechanism throughout. The initial stages of unfolding by heat at neutral pH and by alkali do appear to follow a similar path (Casal *et al.*, 1988)

Alkali Denaturation

Alkali denaturation of bovine BLG becomes significant above pH 8. As the pH rises from 8 to 9 the sedimentation coefficient decreases to a value consistent with a monomeric structure (Townend *et al.*, 1960b), and it remains at this value up to pH 10. Above this, further structural changes occur. ORD studies have shown that around pH 11 the α-helix remains intact whilst the ß-sheet is converted to a random conformation (Timasheff *et al.*, 1966a). In complete contrast, infrared spectroscopy shows that it is the helix which unfolds first in the initial stage of alkai denaturation (Casal *et al.*, 1988). FTIR reveals that at pH 12 the structure is completely random (Purcell & Susi, 1984). These observations were obtained with freshly-made solutions since alkaline denaturation is time-dependent. The rate of denaturation rises rapidly with pH and leads eventually to the formation of large aggregates (Groves *et al.*, 1951; Hin Bon Hoa *et al.*, 1973).

Thermal Denaturation

Although the thermal denaturation of BLG is unlikely to be connected with its function *in vivo*, studies have been carried out which provide information about the environments of specific amino acids. Information about the interactions between BLG and other milk components, which are of interest to industry will be dealt with elsewhere.

Thermal denaturation of bovine BLG *in vitro* has suggested that upon increasing the temperature from 30 to 55°C, the dimer dissociates to monomer (Dupont, 1965; Sawyer, 1969). At higher temperatures, unfolding occurs concomitant with increased activity and oxidation of the thiol group (Larson & Jenness, 1952). Differential scanning calorimetry (DSC) reveals a denaturation peak at about 80°C, pH 6.5, which is pH dependent, and moves to about 60°C at pH 8 (de Wit & Klarenbeek, 1981). Park & Lund (1984),however, observe no such decrease over the same pH range under nearly the same experimental conditions, except that 50mM buffer is present. A shift in the denaturation peak to around 60°C is also observed when the DSC experiment is carried out in the presence of ß-mercaptoethanol. The two Trp residues become exposed, one fully and one partially (Mills, 1976). It is likely that Trp-61, which is located on a flexible external loop, is the one that becomes fully exposed, whilst Trp-19, normally fully

buried at the foot of the rigid ß-barrel becomes only partially exposed. The effect of pH on the conversion of the monomers to a denatured form has indicated that the two imidazoles present per subunit are involved (Dupont, 1965). As His-146 is close to the α-helix, this implies some destruction of the helical structure which is not inconsistent with the large change in optical rotation observed, although there is still some residual secondary structure after heat treatment at 90°C (Ananthanarayanan *et al.*, 1977). This is also consistent with the infrared spectroscopic results which show the initial unfolding of helical structure (Casal *et al.*, 1988). In contrast, lowering the pH has been shown to increase the thermal stability of protein and is suggested that extra hydrogen bonding is responsible (Kella & Kinsella, 1988)

The effect of heating BLG in the presence of other milk components has also been investigated. Studies *in vitro* have shown that lactose stabilizes BLG against thermal denaturation (Park & Lund, 1984), by forming a browning complex which is believed to be antigenic (Otani *et al.*, 1985a) κ-casein on the other hand destabilizes BLG, the enhanced rate of its unfolding being entropy-driven, and indicative of hydrophobic residues becoming exposed (Park & Lund, 1984). The interaction between κ-casein and BLG is believed to involve the free thiol, the disulphide bridges and Ca^{2+} (Sawyer, 1969).

Organic Compounds

BLG can be denatured by compounds like urea, guanidinium hydrochloride, alcohol and detergent. Urea denaturation is believed to occur via a two-step process. At pH 3.5, where BLG tends to be monomeric, the first step is the first-order, reversible unfolding of the polypeptide chain to a random-coiled conformation (Greene & Pace, 1974). However, at pH 5.2, this step follows more complex kinetics which, upon addition of NEM which dissociates dimeric BLG, the kinetics revert to first-order (McKenzie & Ralston, 1971,1973). This indicates that the first stage involves both dissociation and unfolding. The second stage involves the unfolded protein undergoing -SH/SS exchange reactions (and possibly -SH oxidation), to yield irreversible aggregates. Creighton (1980) has used urea-gradient electrophoresis to study the kinetics of the refolding of urea-denatured BLG and Alexander & Pace (1971) have shown that the midpoint concentrations for the urea denaturation of bovine A, B and goat BLG at 25°C and pH 3 are 5.3, 5.19 and 4.68 M. These values show that BLG-A is 1 kJ/mol more stable then BLG-B, and 4kJ/mol more stable than goat.

The effect of ethanol on the structure of bovine BLG has also been examined. Ethanol unfolds the protein at pH 3 in a similar manner to urea (Tanford & De, 1961) although unlike urea, the reaction is not entirely reversible with the unfolded protein refolding to a more α-helical conformation. Timasheff *et al* (1966a) observed a similar increase in helix content in MeOH. It has also been observed that the optical rotation does not appear to vary for concentrations of EtOH less than 10%, although Hemley *et al.* (1979) advise against using more than 5% EtOH when wishing to study the native conformation of BLG in an ethanolic solution.

Sodium dodecyl sulphate (SDS) can also affect the conformation of bovine BLG, but only at high concentrations. BLG binds one mole of the dodecyl sulphate anion per subunit, without any change in the secondary or quaternary structure (Siebles, 1969). However at higher concentrations, an 'all-or-none' complex forms with 22 moles of SDS adsorbed onto the protein surface (Jones & Wilkinson, 1976; Hilquist *et al.*, 1982), but with negligible change to the secondary structure. Continued addition of SDS causes further associations, and a change in secondary structure as the amount of α-helical structure increases. The CD spectrum in the presence of a fifty-fold molar excess does not alter appreciably in the near uv region but the far uv spectrum is considerably altered, especially at low pH (Su & Jirgensons, 1977). Urea denaturation has been used by Cupo & Pace (1983) to monitor the effect of small changes to the environment of the free thiol at pH 2.8; adding $-SCH_2CH_2X$ where X = -COOH, $-CH_3$, -OH, $-NH_2$, progressively destabilises the native structure by amounts varying from 4.5 - 30 kJ mol^{-1} with respect to the native protein.

Heavy Metals

The addition of heavy metal compounds to BLG in solution can induce an irreversible conformational change, prevented if the protein is in crystalline form. The crystalline array clearly stabilizes the native conformation (Dunnill *et al.*, 1966; Green *et al.*, 1979) enabling the isomorphous heavy atom derivatives necessary for X-ray crystallographic studies to be obtained by soaking protein crystals with heavy metal compounds. However, in some cases, even the solid phase, conformational changes appear to be occuring (Sawyer & Green, 1979).

The denaturation of BLG by heavy metals above pH 6 has been studied using mercurial compounds, silver(I) and copper salts. Ag(I) and some Hg(II) salts. For example *p*-chloromercuribenzoate (PCMB) causes an increase in laevorotation, up to a ratio of 2 moles salt per dimer BLG (Pantaloni, 1965) by binding to the free thiol and enhancing dissociation. Hg(II) acetate also shows similar behaviour; although at pH 6.8 the binding of only one mole of Hg(II) per dimer of BLG also shows an increase in laevorotation consistent with a mercurial bridge between the free thiols of each subunit (Lontie & Preaux, 1966). For this to occur, some rearrangement of the protein is necessary as the thiol groups are about 25Å apart in the crystallized native protein (Green *et al.*, 1979). The binding of one Hg(II) per dimer has been studied (Pantaloni, 1964): its reaction is first-order, specific for the free thiol, and involves a Trp although no direct interaction between these residues has been proved. The positions of the Trp residues in the crystal structure are 9 and 31 Å from Cys-121. The denaturation of BLG by Cu(II) occurs by two consecutive, first-order reactions. The first involves the binding of one Cu(II) per subunit (Pantaloni, 1965), possibly at an amide or anomalous carboxyl site, whilst the second stage involves a conformational change of the complex (Pantaloni, 1962). The subsequent loss of copper, and the low concentrations used suggest that it might in some way enhance the Tanford transition.

SECONDARY STRUCTURE

The secondary structure of BLG in solution has been studied by a variety of techniques. Raman spectroscopy showed a strong band at 1242cm^{-1} corresponding to ß-sheet and random coil conformations, and a weak band at 938 cm^{-1} from which it was deduced that the protein possessed about 10% α-helix (Frushour & Koenig, 1975). ORD studies have confirmed an α-helical content of about 10%, and suggested that the remaining 90% was divided evenly between ß-sheet and random coil (Timasheff *et al.*, 1966a); whilst CD experiments gave values for each conformation of 10% α-helix, 43% anti-parallel ß-sheet and 47% unordered (Townend *et al.*, 1967). These values agree with data obtained from IR (Timasheff & Susi, 1966; Rüegg *et al.*, 1975; Casal *et al.*, 1988), FTIR (Byler & Susi, 1986), from some secondary structure prediction algorithms (Deckmyn & Preaux,1978; Creamer *et al.*, 1983), and from the 2.8Å resolution crystal structure, which revealed 51% anti-parallel ß-strands and 7% α-helix (Papiz *et al.*, 1986). Byler *et al.* (1983) have reported a laser-Raman study

of the cystine groups in the 500-700 cm^{-1} spectral region that reveals that both disulphides adopt a *gauche-gauche-gauche* conformation. As has been noted above, extremes of pH and modification of the properties of the solvent alter these values markedly.

CONFORMATIONAL CHANGES

Circular dichroism, and the related technique of optical rotatory dispersion, provide an excellent means of monitoring conformational change. ORD, or even optical rotation, and proton-binding experiments (titration curves) have shown that bovine BLG undergoes three pH- dependent conformational transitions which can be summarized as:

Q <-> N <-> R -> S

FIGURE 1 shows the curves obtained by McKenzie & Sawyer (1967), together with the curve from our laboratory for pig BLG, and the values typical of denatured protein.

The first transition (Q<->N) between pH 4 and 6 is reversible, and has been described for the bovine BLG variants A, B and C (Timasheff *et al.*, 1966b; McKenzie & Sawyer, 1967). The increase in sedimentation coefficient observed, with increasing pH, correlates with the contraction of the protein. BLG-A undergoes a two-group ionization whereas BLG-B follows a single proton transition, suggesting that the additional carboxyl in BLG-A, Asp-64, may be one of the groups being ionized. Also, the pH at which octamers form most readily with BLG-A corresponds to a significant dip in the curve for that variant alone. BLG-C also follows a single proton transition, but unlike BLG-B, one extra cationic residue per subunit, possibly the additional His-59, is exposed upon increasing the pH. From ORD studies Timasheff *et al.* (1966a) report that no aromatic residues are involved in this transition, and Townend *et al.* (1969) also come to this conclusion using solvent perturbation difference spectroscopy. The latter is slightly surprising as His-59, Trp-61 and Asp-64 are located in the same region of the protein. No change is detected over this region in the IR spectrum (Casal *et al.*, 1988)

Between pH 6.5 and 7.8 the second reversible conformational change (N-R), often called the Tanford transition, is observed (Tanford *et al.*, 1959). In bovine BLG this transition can be detected by a change in optical rotation ($[\alpha]_D$ is -25° at pH 6, but -48° at pH 8), by a decrease in the sedimentation coefficient (3.2 to 2.6 S) or possibly even by a change in the position of the thermal denaturation peak observed by DSC (de Wit & Klarenbeek, 1981). The change in sedimentation coefficient may be due to an expansion in the volume of the protein, a variation in the shape of the protein (Tanford *et al.*, 1959; Timasheff *et al.*, 1966b) or the increased dissociation of dimer to monomer (Georges *et al.*, 1962; Pantaloni, 1965; McKenzie & Sawyer, 1967).

Which amino acids are responsible for this conformational change is still unclear, although the effect of this transition on some residues is known. Upon increasing the pH one buried carboxyl per subunit becomes exposed and ionized, with a positive enthalpy which suggests that the buried carboxyl was originally hydrogen-bonded (Tanford & Taggart, 1961). The anomolous carboxyl is a consequence of the native structure since no such observation is made when titrating urea-denatured BLG. The pK_a of this anomalous carboxyl is around 7.3, similar to that of a His. Goat also has an an anomolous carboxyl group (Ghose *et al.*, 1968). Cu(II), which is known to alter the Tanford transition (Pantaloni, 1962,1965), can bind to both carboxyl and imidazole groups, thus making it unclear whether His is involved. Absorption spectroscopy has shown that a Tyr is involved - it is believed that a hindered Tyr becomes partially exposed (Pantaloni, 1965). The free sulphydryl also becomes more accessible when the pH rises from 6 to 8, as demonstrated by an increase in the second order rate constants for the binding of PCMB (Lyster, 1964; Dunnill & Green, 1965) or tetracyanoaurate(III) (Sawyer & Green, 1979) to bovine BLG. The Tanford transition is dramatic when observed by optical rotation but much less so by other techniques. In particular, a large change in the fold in the polypeptide is not revealed, ellipticity change at 222nm (Su & Jirgensens, 1977), or by the crystal structures at pH 6.5 and 7.6. Rather, a subtle rearrangement of side chains is observed (Yewdall, 1988). Further, it is not yet clear that either change is important for the function of the protein although the pH range in which they occur coincides with that in which the protein can be found *in vivo*.

The third, irreversible, conformational change is the alkali denaturation of BLG, which has been discussed previously.

AMINO ACID ENVIRONMENTS

A large number of protein chemical investigations have been carried out to determine the environment of specific amino acids within the protein, many of which were summarised by Townend *et al.* (1969). These studies have mostly been performed with bovine BLG although a few have used the caprid or ovine protein. Within the ruminant proteins, however, the results appear to be broadly similar. Far less is known about the monomeric BLG's.

An obvious target for modification is the free cysteine, shown to be Cys-121 by both crystallographic studies (Papiz *et al.*, 1986; Monaco *et al.*, 1987; Yewdall, 1988) and a series of enzymatic digestion experiments (Mainferme *et al.*, 1971; Martial *et al.*, 1971;Perez *et al.*, 1971). However, the possibility of disulphide interchange between positions Cys-119 and Cys-121 (McKenzie & Shaw, 1972; McKenzie *et al.*, 1972) has not been eliminated. The environment of the free thiol has been investigated using a variety of compounds. The rate of binding of pCMB increases above pH 6 by a factor greater than can be accounted for by the ionization of the -SH, indicating that the thiol becomes more exposed at higher pH (Dunnill & Green, 1965). It is known that when NEM binds to the free sulphydryl, the protein is completely dissociated into subunits (Lontie & Preaux, 1966), implying a location near the site of association. However, dissociation upon sulphydryl modification is by no means invariable (Cunningham & Nuenke, 1960; Townend *et al.*, 1969; Stone & Wishnia, 1978; Cupo & Pace, 1983). Thus, whether the pH-dependent reactivity of the Cys residue results from the N<->R conformational change or dissociation (or, indeed, both) is still unclear. From a comparison of the dimensions of the different substituents Townend *et al.* (1969) suggested that the free thiol and dimer association site must be at least 6Å apart, and that the contact area for octamer formation must be in the same vicinity. Examination of the relative positions of the free thiol and the region of subunit association in the crystal structure shows that this is correct: the helix at the top of the molecule runs between Cys-121 and the molecular dyad axis.

There are four tyrosine residues per subunit of bovine BLG-A and their accessibility has been investigated by reactions with N-acetylimidazole and cyanuric fluoride (Gorbunoff, 1967; Townend *et al.*, 1969). Two of the tyrosines are readily available at pH 9.3 where one can be acetylated with no change in the CD spectrum. The third is partially hindered and only becomes reactive at pH 10.8 after some conformational change of the protein, and the fourth is unavailable, titrating with an apparent pK_a of 12.3 (Tanford *et al.*, 1959). In the case of the goat protein, the reactivity of the third tyrosine is slightly greater. In both cases, these residues are not involved in subunit association, although one or more are affected by the Tanford transition (Pantaloni, 1965).

Bovine BLG contains two tryptophans per monomer, Trp-19 and Trp-61. Both solvent perturbation spectroscopy (Townend *et al.*, 1969) and Trp fluorescence studies (Mills, 1976) have revealed that about half of these residues are buried, but neither distinguished between one buried and one exposed residue, or both partially buried. The structure reveals that Trp-19 is buried and Trp-61 is in an undefined region which is most probably mobile. In contrast, the modification of BLG by 2-hydroxy-5-nitrobenzyl bromide (HNBB) at pH 2 where the protein is a monomer, revealed that both Trp residues were accessible. Further, the presence of retinol had no effect on the Trp modification although the converse is not true (Fugate & Song, 1980). These workers also find that neither Trp is affected over the pH range 2 to 7.5. Townend *et al.* (1969) have excluded the involvement of either Trp residue in the association/dissociation of the dimer or the conformational transition between pH 4 and 6. As electron transfer between Trp and Tyr amino acids is predominantly initiated by the electron-deficient Trp-61, and its rate is independent of pH over the range 6.1 to 7.9, this is in keeping with neither Trp residue being involved in the Tanford transition (Prutz *et al.*, 1980).

Phosphorylation of BLG by $POCl_3$ in CCl_4 produces no phosphoserines, the major species being N-phospholysine and/or N-phosphohistidine (Woo *et al.*, 1982). The extent of phosphorylation increases with pH, but never exceeds a stoichiometry of 14 mol phosphate/mol BLG. This is less than the total number of Lys and His sites (15+2) indicating that some are not fully exposed. Most of the Lys residues in bovine BLG are accessible for reductive alkylation although those in the α-helix (Lys-135,-138,-141) are only partially modified, and Lys-47 remains unaltered (Brown *et al.*, 1988). Cleavage of the penultimate His and C-terminal Ile of bovine BLGs has been carried out using carboxypeptidase A (Greenberg & Kalan, 1965).

The carboxyl groups in BLG have also been studied. Treatment of bovine BLG with carbodiimide has shown that modifying 7-8 of the carboxyls prevents octamerization of BLG-A (Armstrong & McKenzie, 1967). Esterification with MeOH, EtOH and BuOH affects 23, 19 and 12 carboxyls respectively out of the 25 per subunit (Halpin & Richardson,1985). Of the two hindered carboxyls not modified by MeOH, one is likely to be the anomalous carboxyl associated with the Tanford transition. Mattarella & Richardson (1983) have shown that the carboxyl modification leads to significant changes in the character of the protein. A detailed comparison of the crystal structures at either side of this transition will enable the anomalous carboxyl to be identified unambigously.

BINDING STUDIES.

A wide variety of compounds is known to bind to BLG, some of which are listed together with their association constants in **TABLE IX**. However, it is still unclear which regions of the protein are involved, and whether any of the ligands which bind are important for the biological function of this protein.

Aliphatic Molecules

Probably the first binding was shown by the surface area measurements of Bull (1946b) who found 8 molecules of sodium lauryl sulphate (SLS) associated with the BLG dimer. Since continued addition of SLS led to a sharp change at 40 mol/dimer, Bull considered the effect to be one of titration rather than association. It is not surprising, therefore, that SDS has since been shown to bind to bovine BLG. The first stage involves the tight binding of two anions/dimer to yield a complex which has been crystallized (McMeekin *et al.*, 1949), but not yet studied by X-ray crystallography. Chemical modifications of BLG before complexation revealed that neither Trp, Cys nor His-161 were essential for binding, but did show that the hindered His, His-146, was important for the formation of the BLG+SDS complex (Seibles, 1969). Higher concentrations of SDS caused a conformational change associated with the binding of 22 anions. This complex is in equilibrium with that to which 2 SDS anions are bound (Hillquist *et al.*, 1982; see above).

Free fatty acids (FFA) can also bind to bovine BLG in a manner similar to SDS. They reversibly bind at a single strong binding site per subunit with an association constant of about $10^5 M^{-1}$, inducing a conformational change. A further binding at up to 24 weaker secondary sites presumably similar to those for SDS, has been reported (Spector & Fletcher, 1970). Binding at the primary site depends upon the nature of the ligand, decreasing in the order palmitic(C16:0) > stearic(C18:0) > oleic(C18:1) > lauric (C12:0). Further studies using palmitate and bovine BLG have indicated that the binding affinity increases over the pH range 6.5 to 8.5, suggesting a possible correlation with the Tanford transition. Spin-labelled stearate derivatives *eg* 12-doxyl stearic acid have been studied by ESR and show some weak binding to BLG, but do not form 1:1 complexes (E.O. Keith, personal communication, 1987). This suggests that the alkyl chain binding site has little spare capacity. Interestingly, the FFAs with C > 14 which bind to BLG, are those most prevalent in milk. Whether this is coincidence, or because BLG has a role as a general FFA carrier in the lipid metabolism of the mammary gland is still unclear (Diaz de Villegas *et al.*, 1987).

Triglycerides bind to BLG (Smith *et al.*, 1983), but no details of the binding are given. Phosphatidylcholines appear unable to bind to native BLG (Brown *et al.*, 1983) but in the presence of a helix-inducing solvent, BLG can refold to a structure containing about 50% α-helix, which is then able to bind dipalmitoyl phosphatidylcholine. The vesicles formed from the sonication of the complex revealed that the ratio of lipid:BLG was 20:1, in keeping with both the SDS and FFA complexes. Further, CD revealed that BLG now contained 30% α-helix. BLG in this form also exhibited a 10% increase in Trp fluorescence, compared to 8% when a FFA binds, although the hydrophobicity of the residues remained unaltered, indicated by a reduction in quenching (Brown *et al.*, 1983). The ability of partially unfolded BLG to bind to lipids may be of importance for its transport across a membrane. More recently Cornell and Patterson (1989) showed that BLG can adsorb to an artificial phospholipid membrane surface at pH 4.4 though not at pH 7.0. CD measurements showed no change in conformation on adsorption so that an electrostatic mechanism has been proposed. This is consistent with the large dipole moment of the molecule and also with the adsorption to colloidal gold (Horisberger & Vauthey, 1984).

Many alkanes bind to BLG, the amount bound decreasing as the chain length increases (Mohammadzadeh-K. *et al.*, 1969). Butane, pentane and iodobutane can all bind to a single hydrophobic region per subunit of bovine BLG (Wishnia & Pinder, 1966). The site, believed to be a hydrophobic pocket going from the surface towards the interior of the protein, is unaffected by the state of association of the protein. This pocket can bind either two butanes equally, two pentanes unequally, possibly for steric reasons or because of exposure of a hydrocarbon tail, or one iodobutane. The position of the latter, soaked into BLG crystals and determined by X-ray crystallography, locates this alkane-binding site inside the hydrophobic cavity (Papiz, 1982).

Aliphatic ketones *eg.* 2-heptanone, also bind at a single hydrophobic site in each subunit of bovine BLG, with association constants of the order of $10^3 M^{-1}$. As these are of the same

magnitude as those for alkane binding, these ketones are probably occupying the alkane binding cavity (O'Neil & Kinsella, 1987).

Conjugated Compounds

Both fused and single ring aromatic compounds are known to bind to BLG. The binding of N-methyl-2-anilino-6-naphthalenesulphonate at pH 8 has been studied by calorimetry, equilibrium dialysis and fluorescence spectrometry, and reveals that bovine BLG-B possesses two anionic binding sites, one of which binds more strongly than the other (Lovrien & Anderson, 1969). Both of these sites are hydrophobic and one is likely to correspond to that above, although as has been seen the alkyl binding site is apparently restricted in its capacity. The rather feeble binding to a single site at pH 6 found by Lovrien & Anderson is probably that for 8-anilino-naphthalene-1-sulphonate examined by Mills & Creamer (1975) between pH 2 and 6.5.

The hydrolysis of *p*-nitrophenyl phosphate (PNPP) by bovine spleen phosphoprotein phosphatase was partially inhibited by BLG, suggesting that BLG could bind PNPP (Thompson & Farrell, 1974). The binding of several *p*-nitrophenyl compounds to bovine BLG-A was confirmed by Trp fluorescence studies (Farrell *et al.*, 1987), implicating Trp in the binding. In addition, the dissociation constant for the 1:1 complex of bovine BLG and PNPP was independent of pH over the range 4 to 7.5, indicating little dependence on conformational state.

The binding of toluene, trifluorotoluene and hexafluorobenzene to bovine BLG has been investigated, and reveals two distinct forms of binding : strong association at a single hydrophobic site per subunit with an accessible volume of around 220 ml/mol, which is believed to be the alkane binding site, and weaker interactions at one or more other sites. Neither type of association is affected by dimerization of the protein (Robillard & Wishnia, 1972a,b) although the octamerisation of BLG-A is abolished.

By what may have been a chance experiment, Futterman and Heller (1972) showed that BLG enhanced the fluorescence lifetime of retinol from 3 to 10 ns and reduced its susceptability to oxidation. UV absorption spectroscopy has shown that BLG and retinol form a complex (Hemley *et al.*, 1979), whilst CD has revealed that one retinol is bound per subunit, with a dissociation constant of 0.02μM (Fugate & Song, 1980) compared with 0.2μM for the retinol-RBP complex. The location and nature of the retinol binding site in BLG is uncertain. Modification of the Trp residues with 2-hydroxy-5-nitrobenzyl bromide (HNBB) caused a blue shift in the absorption spectrum of the BLG-retinol complex to that of free retinol (Fugate & Song, 1980), implying that bonding interactions have been destroyed, and implicating a Trp in the binding site. Surprisingly, retinol still binds to the protein in 8M urea (or SDS). The binding is independent of pH in the range 2 to 7.5, suggesting that either a non-ionizable group is involved or the binding site is hydrophobic. Since HNBB treatment can also affect free sulphydryls, involvement of Cys-121 has not been ruled out. Crystallographic studies of the trigonal lattice Z, grown at pH 7.8 from a solution containing bovine BLG and excess retinol have shown that retinol apparently binds in an external hydrophobic channel, adjacent to the α-helix and Cys-121, and not in the hydrophobic cavity near Trp-19 (Monaco *et al.*, 1987).

Ions

There is evidence for the binding of both monovalent and divalent cations to bovine BLG (Baumy & Brule, 1988). Na^+ is able to bind to carboxyls and imidazoles in this protein above its pI value, the number of ions bound per subunit increasing from 0 to 2 as the pH rises from 6 to 9 (Baker & Saroff, 1965). The binding curve is of a similar shape to that observed by optical rotation over the Tanford transition, although slightly displaced to higher pH. Various heavy metal cations have also been shown to bind: Ag^+ (Pantaloni, 1965), Cu^{++} (Pantaloni, 1962) and mercurial salts (Pantaloni, 1964), and these can form complexes with one cation bound per free sulphydryl. However, Hg(II) acetate above pH 6.8 can also form a complex in which there is only one cation bound per dimer of BLG. It is believed that the Hg^{++} ion forms a bridge between the free thiol in each subunit (Lontie & Preaux, 1966), although for this to be the case considerable rearrangement and/or disulphide interchange of the protein must occur since the free thiols are about 25Å apart in the native structure (Green *et al.*, 1979).

One type of anion which is known to bind to BLG the Iodide ion. I_3^- reacts specifically with the free thiols of bovine BLG producing a sulphenyl iodide complex. This is less stable than the native protein and is able to react with different mercaptans to yield derivatives with varying abilities to octamerize (Cunningham & Nuenke, 1960). Wishnia & Stone (1978) have shown that there is a single site for binding HgI_3^- on BLG when the free thiol group is blocked with mercaptoethanol. This site most probably is that identified by Papiz *et al.* (1986) as the major site for the "$HgI_4^{=}$" derivative, close to the molecular dyad axis. This site is not that for alkane binding either, though when native BLG is used, the binding of mercuriiodide complexes in solution is less straightforward.

These binding studies indicate that BLG has the ability to bind hydrophobic or amphipathic molecules to at least two distinct sites. Neither site interferes with or is affected by the monomer or dimer state of the protein, but higher association is interrupted by one of them. Binding at one site involves changes in the sulphydryl environment which may lead to dissociation whereas binding at the other site appears apears in some cases to involve a tryptophan residue. Surprisingly, denaturation does not apear to abolish the binding of some ligands. Examination of the native structures reveals a deep, calyx-like cavity which by analogy with the homologous retinol-binding protein, is able to accommodate retinol but which is quite restricted in its lateral dimensions. In both structures, a Trp residue is at the foot of the calyx. Further, the horse-shoe shaped bilin molecule binds at the entrance to this cavity in insecticyanin (Holden *et al.*, 1987) and bilin-binding protein (Huber *et al.*, 1987), both of whose structures show the same polypeptide fold as BLG. Tempting though it is to claim that this internal site is that for vitamin A, electron density not inconsistent with vitamin A has been found on the outer surface of the calyx, close to the helix, in one of the high pH crystal forms co-crystallised with the ligand (Monaco *et al.*, 1987). This site also appears to be hydrophobic in nature and is less restricted since it is a crevice on the outer surface of the molecule. Crystal soaking experiments are hampered by the instability of the crystals in solvents which can dissolve ligands like vitamin A or alkanes so that it has proved difficult to obtain convincing direct evidence on the detailed nature of the binding sites.

MACROMOLECULE BINDING

Non-milk Proteins

Cytochrome-c has been reported to interact with BLG. The initial binding to subunits of bovine BLG produces a 1:1 complex. If the pH is above 7.5, the conformation of BLG in the complex is altered, thus enabling the free sulphydryl of BLG to reduce the iron in cytochrome-c (Brown & Farrell, 1978). A non-reduced complex has also been observed between cytochrome-c and pig

BLG since the latter has no free thiol. It is unclear at present if it is the N<->R conformational change which leads to this interaction, or which amino acids are involved. BLGs A,B,C all form a similar complex with cytochrome-c, although the rate of iron reduction depends on the variant. Other proteins have also been shown to associate with BLG, mostly through inhibition of enzyme activity. For instance, the inhibition of human renin by bovine BLG-B has been reported (Workman *et al.*, 1974). It is non-competitive, and requires a leucine residue near a hydrophobic region in BLG. Similarly, Farrell & Thompson (1971) have investigated the effects of BLG on phenylphosphate hydrolysis by bovine spleen phosphoprotein phosphatase as mentioned above.

Milk Proteins

There are numerous studies of the interactions between milk proteins, many resulting from the milk-processing and food industries, which are outwith the scope of this review. Bovine BLG is known to interact with α-lactalbumin (Hunziker & Tarassuk, 1965), and several of the caseins. For example, it can react with κ-casein to form a 3:1 complex involving hydrophobic interactions. This complex is subsequently stabilized by covalent bonding, and a conformational change which makes the disulphide bridges less susceptible to attack (Haque *et al.*, 1987), possibly therefore as a result of disulphide interchange.

Immunoglobulins

Antibodies raised against bovine BLG-B cross-react with ruminant BLGs to different extents. This appears to be dependent upon the immunochemical method used to detect the cross-reactivity as well as the species from which the BLG originates (Jenness *et al.*, 1967). It is therefore possible that the few amino acid substitutions between ruminant BLGs occur in the antigenic regions of the protein. Cross reaction with antisera against bovine BLG and monomeric BLG from pig, horse (Liberatori *et al.*, 1979a) or camel (Liberatori *et al.*, 1979b) is low, indicating some similarity in the two types of protein, though with the possibilty that the antigenic region(s) involve both subunits.

Antibodies towards bovine BLG have been shown to cross react with human milk, although the presence of bovine BLG in human milk is most likely to be a transfer of the protein or at least antigenic peptides, from the mother's diet (Axelsson *et al.*, 1986). Indeed, on removal of milk proteins from the mothers diet the levels fall to non-detectable amounts (Jakobsson *et al.*, 1985). Similarly, antibody studies have revealed the presence of bovine BLG in cow's urine after cessation of milking for prolonged periods (Lyster & Wheelock,1967) and in the urine of new-born suckled calves (Pierce, 1960).

As well as using antibodies to bovine BLG to detect the presence of BLG in the milks and other fluids, they are also useful in monitoring the presence of BLG in the gastro-intestinal (GI) tract, and for identifying the allergenic and antigenic sites of BLG.

Koritz *et al.* (1987) used antibodies towards bovine BLG, raised within a rabbit, to show that the concentration of BLG decreases upon its progression down the GI tract owing to its increased exposure to proteinases and intestinal absorption. Levels remain high enough to detect up to two days after removal of BLG from the diet, after which the levels drop (Suzuki *et al.*, 1987). These antibodies were tentatively identified as secretory IgA (sIgA) and have also been detected in low levels in the saliva of infants, although their origin is unclear (Frick & Rieger, 1987).

Increasing the amount of cows milk ingested has been shown to cause an increase in the amount of IgG in serum (Fallstrom *et al.*, 1978). Challenges with isolated milk proteins confirm that BLG is often the offending agent in infants with an allergy to cows milk (Lebenthal *et al.*, 1970). It is thought that the immature nature of the neonate intestinal surface enables BLG to be absorbed, raising IgG levels and complexing with them (Robertson *et al.*, 1982). Why only some infants are susceptible is unclear, although the nature of the ingested BLG may be important. Huang *et al.* (1985) have investigated tryptic and cyanogen bromide fragments for both allergenicity and antigenicity. A monoclonal IgE towards bovine BLG showed reactivity with both native and aggregated forms of the protein, however only the latter induced positive cutaneous anaphylaxis (Granato & Piquet, 1986).

Ingestion by guinea-pigs of bovine BLGs A and B leads to higher levels of antibody being raised towards the A variant. However, injection into sensitised guinea-pigs induces anaphylactic shock when challenged with either variant, indicating the same allergenic site in both variants distinct from the antigenic sites which bind to IgG. It is possible that the antigenic site involves one or both of the two substituted amino acids, Asp-64 and Val-118, in the A variant (Malik *et al.*, 1988). IgE and IgG antibody binding, to reduced and carboxymethylated BLG, indicates that the disulphide bridges are unimportant in allergenicity, though important in the antigenic region (Huang *et al.*, 1985).

Various fragments and derivatives of bovine BLG have been exposed to IgE and IgG antibodies raised against native BLG, to try to locate the antigenic sites. Tryptic, chymotryptic and peptic digests of BLG show no cross reaction with IgE raised towards native protein suggesting that the antigenic sites are discontinuous (Kurisaki *et al.*, 1982). Kurisaki *et al.* also indicated the importance of the disulphide bridges but noted that modifications of the free thiol, either Trp, two of the three Arg residues or the amino groups did not inhibit cross reaction on incubation with IgE. Four antigenic sites have been proposed for native BLG which are present at both pH 6 and 8.5. This indicates that neither the Tanford transition, nor the monomer-dimer state of association significantly effects the antigenic nature of the protein (Kurisaki *et al.*, 1985). Identical modifications to those in their 1982 paper gave similar results with IgG antibodies except that Trp and amino group modification did affect the IgG interaction. This may indicate the importance of these sites in antigenicity but not allergenicity. Additional modifications have also implicated histidine, carboxyl and possibly tyrosine sidechains in the antigenic sites. However, CD measurements on chemically modified BLGs indicate that the structures may no longer be in their native conformation (Otani *et al.*, 1985b).

Some recent work may give a better indication of the position of an antigenic region within the bovine BLG. Monti *et al.* (1989) showed a reaction between antibodies towards bovine BLG and milk from humans fed on a milk free diet. Recognition was attributed to a fragment at the N-terminus of human lactoferrin. Note that the cross-reactivity of lactoferrin with anti-bovine BLG antibody is an alternative explanation for the reports of a human BLG. Further, sequence comparisons between β_2-microglobulin and BLG, initiated because of cross reactivity with some antibodies to BLG, implicate a region between residues 121 and 140 in BLG that are responsible (Conti & Godovac-Zimmerman, 1990). This is identical to that proposed by Monti *et al.* (1989) and corresponds to the main stretch of α-helix.

CRYSTALLOGRAPHIC STUDIES

Many of the methods used to isolate BLG from milk produce an impure protein, which can be purified by recrystallization. Microcrystals of BLG have been obtained during the preparations of BLG from the milks of ruminants: cow (Palmer *et al.*, 1934; Aschaffenburg & Drewry, 1957), goat (Askonas, 1954), and sheep (Maubois *et al.*, 1965), and of non- ruminants: pig (Kalan *et al.*, 1971). These were often too small or too disordered to be suitable for X-ray crystallography but crystals suitable for this type of structural study have been grown from bovine and buffalo BLGs, and characterized (Crowfoot & Riley, 1938; Senti & Warner, 1948; Green *et al.*, 1956; Aschaffenburg *et al.*, 1965; Bolognesi *et al.*, 1979). These papers reveal that there are at least 12 crystal forms of the protein grown under various conditions of pH, salt concentration and the presence of ligands. Of these, the salted-out forms have proved to be the most satisfactory for X-ray crystallography.

BLG was one of the first proteins to be subjected to X-ray analysis. Work on both wet and air-dried crystals grown without salt, revealed the dimeric nature of the protein although this was not recognised at the time (Crowfoot & Riley, 1939). The dimeric nature of the protein in these and other crystal forms was discussed by Green and Aschaffenburg (1959) and an account of the early X-ray work is given by Hodgkin & Riley (1968).

The structures of four distinct crystal forms, lattices K, X, Y and Z have been obtained by X-ray crystallography using isomorphous the method of isomorphous replacement (Green *et al.*, 1979). All four structures show features suggestive of an α-helix and some ß-sheet, but no distinct differences are evident. Unlike lattices X, Y and Z, lattice K was produced from protein which had the C-terminal -His-Ile residues removed with carboxypeptidase. However, at this low resolution, no indication of the position of the C-terminus could be obtained. A comparison of the positions of the three heavy-atom sites in common amongst the lattices X, Y and Z indicated a possible movement of the free sulphydryl group in keeping with the expected conformational transition between pH 6.5 (lattice X) and 7.8 (lattices Y and Z).

Detailed X-ray crystallographic studies have been carried out on three crystal forms of bovine BLG-A, which were obtained by salting-out with ammonium sulphate:

			a/Å	b/Å	c/Å	α	β	γ	
Lattice X	Triclinic	P1	37.8	49.6	56.6	123.42°	97.28°	103.66°	pH 6.5
Lattice Y	Orthorhombic	$B22_12$	55.7	67.2	81.7	90.0°	90.0°	90.0°	pH 7.6
Lattice Z	Trigonal	$P3_221$	54.4	54.4	113.1	90.0°	90.0°	120.0°	pH 7.6

Of the two higher pH forms lattice Z crystals usually appear first, and then convert to lattice Y, providing that the protein concentration is not too high (Papiz, 1982). Each of these forms has so far given rise to an interpretation of the electron density map at 2.8Å resolution or better (Papiz *et al.*, 1986; Monaco *et al.*, 1987; Yewdell, 1988). Work on lattices X and Y is being extended to 2.0 and 1.8Å resolution respectively. As mentioned above, the structure of the lattice Z protein has also been obtained with retinol bound (Monaco *et al*., 1987)

The first structure of bovine BLG to be obtained was that of the orthorhombic lattice Y at 2.8Å resolution (Papiz *et al.*, 1986). The structure, which is shown in **Figure 2**, consists of nine strands of anti-parallel ß-sheet, eight of which wrap round to create a flattened, conical barrel, or calyx, closed at one end. There is a 3 turn α-helix on the outer surface of the calyx. The positions of two important residues, Trp-19 and Cys-121, and the interface between subunits, have been located. Trp-19 is situated at the foot of the ß-barrel, on the inner surface - modelling retinol into the cavity revealed that the ß-ionone ring is 10Å away from the Trp, whilst its hydroxyl is near Lys-70. The free thiol, Cys-121, is located near the surface, in a hydrophobic channel, bordered on one side by the α-helix. The interface between subunits involves hydrophobic interactions between Ile-29, Ile-147 and the stacking of the His-146, and their two-fold related equivalents. Unfortunately, there is a considerable amount of disorder in the larger, outer loops which means that it has not been possible to locate all of the side chains with certainty. However, it is tempting to equate Tyr-102 with the buried one, Tyr-20 with that part buried and associated with the Tanford transition, and Trp-19 as that mostly buried (Townend *et al.*, 1969). The anomalous carboxyl group is harder to assign since several could be considered as being in atypical environments.

BLG and serum RBP show only low amino acid similarity (about 25% identity) (Godovac-Zimmerman *et al.*, 1985a; Pervaiz & Brew, 1985), but tertiary structures show a high degree of similarity (Sawyer *et al.*, 1985; Papiz *et al.*, 1986). The eight-stranded ß-barrel, α-helix and disulphide bridges are conserved in both proteins, with the insertions in RBP relative to BLG appearing external loops. The structure of the isolated RBP-retinol complex revealed that the chromophore was located within the hydrophobic cavity and prevented from penetration to the foot of the calyx by a ring of five Phe residues (Newcomer *et al.*, 1984). Removal of the chromophore produced apo-RBP, for which a crystal structure has not yet been obtained, although its structure has been modelled (Åqvist *et al.*, 1986). RBP *in vitro* can bind a variety of retinoids (Horwitz & Heller, 1974), suggesting that the binding site is not an exact match to retinol. An external binding site has been proposed for the binding of ß-ionylidene ethylbromoacetate to RBP which is located near the carboxy-end of the α-helix (Gawinowicz & Goodman, 1985). This position appears to be similar to the hydrophobic channel to which retinol binds when cocrystallied with BLG (Monaco *et al.*, 1987).

The same ß-barrel with 8-strands and an α-helix near their C-terminus is found in two insect proteins which bind bilin or biliverdin, Insecticyanin (iCN) from *Manduca sexta L.*, and bilin binding protein (BBP) from *Pieris brassicae*. Both the ICN structure at 2.6Å resolution (Holden *et al.*, 1987) and the BBP structure at 2.0Å (Huber *et al.*, 1987) are similar to BLG, RBP and one another. They bind biliverdin IXg (in a cyclic conformation) at the same position within, but near the top of, the hydrophobic cavity where retinol is located in RBP. The orientation of the chromophore differs between these two biliproteins. Despite the structural homology between ICN, BBP and BLG, it is unlikely that BLG binds biliverdin IXg *in vivo*. Although biliverdin is found in buffalos milk, it is only present at low concentration (0.6 µg/ml), is associated with the casein fraction, and exists in a different isomeric form (Kamur *et al.*, 1984).

HOMOLOGOUS PROTEINS

There has emerged over the past decade a series of proteins whose structures are probably closely related to that of BLG. This family has been called the α_{2u}-globulin superfamily. Sawyer (1987) pointed out that the necessary requirements for this family were a short sequence near the N terminus which contained the residues -Gly-X-Trp-Tyr-, another stretch -Thr-Asp-Tyr-, a conserved basic residue and a disulphide bridge between the C-terminus and a residue around 60. These requirements have been further refined (North, 1989a,b) to produce a consensus pattern for the two short stretches of sequence.

-h-(D,N,E,S,T)-X-X-u-(L,I,V,F,Y)-X-**G**-X-**W**-(W,Y,F,H,R,K)-X-(L,I,V,M)-h-

-h-h-X-**T**-**D**-Y-X-X-y-h-

where h is hydrophobic, u is often basic and y is aromatic.

some 30 amino-acids shorter than BLG (Jones *et al.*, 1988b; Sacchettini *et al.*, 1989; McRee *et al.*, 1989). The consensus pattern at the N-terminus is

G-X-(W,Y,F)-X-(I,V,L,M)-X-X-X-N-(F,Y)-(D,E)-,

the first few residues of which are reminiscent of the first pattern in the BLG family.

CONCLUSION

From the above, it can be seen that BLG is a molecule about which much is known. It is capable of binding small molecules and of interacting with large ones, including itself. Now that the protein is available in over-expressed forms in both yeast and *E. coli*, modifications which can affect these interactions should be fairly straightforward since they can be based upon the three-dimensional structure. Such changes to the protein will provide an excellent opportunity to explain fully the conformational, binding and self-association properties of the protein and may even lead to a proper description of its biological function.

Although these common amino acids represent only a small proportion of the sequence, North (1989a,b) has found them to be grouped spatially at the base of the calyx furthest from its entrance. He further proposes a receptor recognition function for this clustering which implies that there are subtle modifications, analogous to those found in antibodies, between the members allowing their quite distinct interaction with their receptors.

TABLE X lists the various members of this family, together with their functions where these are known and the molecules which are known to bind to them. All of the members have molecular weights around 18 kDa, and appear to bind and transport a hydrophobic/labile/insoluble ligand. The relatedness of these proteins is further supported by the gene homology between five members of the superfamily (Ali & Clark, 1988).

FUNCTION

To date, no biological function for BLG has been discovered. The amino acid composition is such that the protein is of high nutritional value but the molecular properties, particularly the acid stability, lead to the supposition that some other, more specific function exists. It might be supposed, too, that some telling correlation between the presence of the protein in the milk of a species and some specific physiological or biochemical function would help direct research towards the true function. However, as yet no such compelling correlation has been derived. Several have been proposed without much evidence (or conviction!) by Thompson & Farrell (1974)including an involvement with phosphorylation-dephosphorylation in the mammary gland or as a transferase. Interestingly, the protein in the lipocalycin family whose sequence is closest is the human placental protein 14, or pregnancy-associated endometrial glycoprotein (PEG). On the other hand, it is unlikely that the same requirement for a particular developmental function in the first trimester of human pregnancy can be met from an ingested milk protein in the neonate calf. However, the similarity of BLG to the wider family of transport proteins, together with its resistance to low pH and gastric proteolysis (Miranda & Pelissier, 1983; McAlpine and Sawyer, 1990), point towards a transport or transfer function in the intestinal tract where specific receptors have been detected in the neonate calf (Papiz *et al.*, 1986). Further, Said *et al.* (1989) have shown that BLG enhances retinol uptake in the jejunum and ileum of suckling rats but whether retinol, fatty acid or some other ligand is the primary substrate is not yet known. Indeed, the question as to whether BLG is loaded with retinol or something else in the colestrum or milk has not yet been satisfactorily answered although indications have been obtained that both retinol (Garrick & Williamson, 1986) and fatty acid (Perez *et al.*, 1989) are bound to the protein in milk.

The disappearance of specific receptors fairly early on in the life of the neonate calf is certainly consistent with the transfer of something derived exclusively from the colostrum or possibly the milk. While a transport function can be proposed on the basis of circumstantial evidence and a few experiments, an involvement with the transfer of passive immunity as has been proposed by Jenness (1982) must remain a possibility. Superficially, the protein fold is not unlike an immunoglobulin domain (both proteins are mostly ß-sheet), but a detailed examination shows them to be quite distinct.

It can be pointed out, too, that the protein family which includes the cellular RBP, p2 myelin, photoreactive yellow protein and fatty acid binding protein appears to be distinct from that of BLG although it is based upon a 10-stranded ß-barrel formed from a polypeptide chain which is

Åkerstrom, B. & Logdberg, L. : (1990) An intriguing member of the lipocalin family: α_1-microglobulin. *TIBS* **15** 240-243

Alexander, & Pace,C.N. : (1971) A comparison of the denaturation of bovine ß-lactoglobulin's A and B, and goat ß-lactoglobulin. *Biochem* **10** 2738-2743

Ali,S. : (1989) PhD Thesis, University of Edinburgh

Ali,S. & Clark,A.J. : (1988) Characterization of the gene encoding ovine ß-lactoglobulin. Similarity to the genes for retinol binding protein and other secretory proteins. *J. Mol. Biol.* **199** 415-426

Ananthanarayanan,V.S., Ahmad,F. & Bigelow,C.C. : (1977) The denaturation of ß-lactoglobulin-A at pH 2. *Biochim. Biophys. Acta* **492** 194-203

Anderson, L.J.,Hayes,G., Pearse,M.J., Stewart,A.F., Willis,T.M. McKinlay,A.G.: (1989) Complete sequence of bovine ß-lactoglobulin cDNA. *Nucleic Acid Research* **17**, 6739

Åqvist,J., Sandblom,P., Jones,T.A., Newcomer,M.E., van,Gunsteren, W.F. & Tapia,O. : (1986) Molecular dynamics simulations of the holo and apo forms of retinol binding protein. Structural and dynamical changes induced by retinol removal. *J. Mol. Biol.* **192** 593-604

Arakawa,T. & Timasheff, S.N. : (1987) Abnormal sòlubility behaviour of ß-lactoglobulin: salting-in by glycine and NaCl. *Biochem.* **26** 5147-5153

Armstrong,J.McD. & McKenzie,H.A. : (1967) A method for modification of carboxyl groups in proteins : its application to the association of bovine ß-lactoglobulin A. *Biochim. Biophys. Acta* **147** 93-99

Armstrong,J.McD., McKenzie,H.A. & Sawyer,W.H. : (1967) On the fractionation of ß-lactoglobulin and α-lactalbumin. *Biochim. Biophys. Acta* **147** 60-72

Aschaffenburg, R. : (1964) Protein phenotyping by direct polyacrylamide-gel electrophoresis of whole milk *Biochim.Biophys. Acta* **82**, 188-191

Aschaffenburg, R. : (1968) Genetic varaiants of milk proteins: their breed distribution *J.Dairy Res.* **35**, 447-460

Aschaffenburg,R. & Drewry,J. : (1955) Occurrence of different ß-lactoglobulins in cow's milk *Nature* **176** 218-219

Aschaffenburg,R. & Drewry,J. : (1957) Improved method for the preparation of crystalline ß-lactoglobulin and α-lactalbumin from cow's milk. *Biochem. J.* **65** 273-277

Aschaffenburg,R., Green,D.W. & Simmons,R.M. : (1965) Crystal forms of ß-lactoglobulin. *J. Mol. Biol.* **13** 194-201

Aschaffenburg, R. & Thyman, M. : (1965) Simultaneous phenotyping process for the principal proteins of cows' milk. *J. Dairy Sci.* **46** 1524-1526

Askonas,B.A. : (1954) Crystallization of goat ß-lactoglobulin. *Biochem. J.* **58** 332-336

Axelsson,I., Jakobsson,I., Lindberg,T. & Benediktsson,B. : (1986) Bovine ß-lactoglobulin in the human milk. *Acta Paed. Scand.* **75** 702-707

Bain, J.A. & Deutsch, H.F. : (1948) Studies on lactoglobulins. *Arch. Biochem* **16** 221-229

Baker,H.P. & Saroff,H.A. : (1965) Binding of sodium ions to ß-lactoglobulin. *Biochemistry* **4** 1670-1677

Baumy,J.J. & Brule,G. : (1988) Binding of bivalent cations to α-lactalbumin and ß-lactoglobulin: effect of pH and ionic strength. *Le Lait* **68** 33-48

Bell,K. & McKenzie,H.A. : (1964) ß-lactoglobulins. *Nature* **204** 1275-1279

Bell,K. & McKenzie,H.A. : (1967a) The isolation and properties of bovine ß-lactoglobulin C. *Biochim. Biophys. Acta* **147** 109-122

Bell,K. & McKenzie,H.A. : (1967b) The whey proteins of ovine milk: ß-lactoglobulins A and B. *Biochim. Biophys. Acta* **147** 123-134

Bell,K., McKenzie,H.A., Murphy,W.H. & Shaw,D.C. : (1970) ß-lactoglobulin droughtmaster : a unique protein variant. *Biochem. Biophys. Acta* **214** 427-436

Bell,K., McKenzie,H.A., Muller,V., Rogers,C. & Shaw,D.C. : (1981a) Equine whey proteins. *Comp. Biochem. Physiol.* **68B** 225-236

Bell,K., McKenzie,H.A. & Shaw,D.C. : (1981b) Bovine ß-lactoglobulin E, F and G of Bali (Banteng) Cattle, *Bos (Bibos) javanicus. Aust. J. Biol. Sci.* **34** 133-147

Bell,K., McKenzie,H.A. & Shaw,D.C. : (1981c) Porcine ß-lactoglobulin A and C. *Mol. Cell. Biochem.* **35** 103-111

Bolognesi,M., Liberatori,J., Oberti,R. & Ungaretti,L. : (1979) Preliminary crystallographic data on buffalo ß-lactoglobulin. *J. Mol. Biol.* **131** 411-413

Braunitzer,G., Chen,R., Schrank,B. & Stangl,A. : (1972) Die sequenzanalyse des ß-lactoglobulins. *Hoppe-Seyler's Z. Physiol. Chem.* **354** 867-878

Braunitzer,G., Liberatori,J. & Kolde,H.-J. : (1979) The primary structure of the ß-lactoglobulin of the waterbuffalo (*Bubalus arnee*). *Z. Naturforsch.* **34c** 880-881

Brew,K. & Campbell,P.N. : (1967) The characterization of the whey proteins of guinea-pig milk. *Biochem. J.* **102** 258-264

Brignon,G., Chtourou,A. & Ribadeau-Dumas,B. : (1985) Does ß-lactoglobulin occur in human milk? *J. Dairy Res.* **52** 249-254

Brignon,G. & Ribadeau-Damas,B. : (1973) Localisation dans la chaine peptidique de la ß-lactoglobuline bovine de la substitution Glu/Gln differenciant les varients genetiques B et D. *FEBS Lett.* **33** 73-76

Brignon,G., Ribadeau-Dumas,B., Garnier,J., Pantaloni,D., Guinard,S., Basch,J.J. & Timasheff,S.N. : (1969) Chemical and physico- chemical characterization of genetic variant D of bovine ß-lactoglobulin D. *Arch. Biochem. Biophys.* **129** 720-727

Brooks,D.E., Means,A.R., Wright,E.J., Singh,S.P. & Tiver,K.K. : (1986) Molecular cloning of the cDNA for two major androgen-dependent secretory proteins of 18.5 kilodaltons synthesised by the rat epididymis. *J. Biol. Chem.* **261** 4956-4961

Brown,E.M., Carroll,R.J., Pfeffer,P.E. & Sampugna,J. : (1983) Complex formation in sonicated mixtures of ß-lactoglobulin and phosphatidylcholine. *Lipids* **18** 111-118

Brown,E.M. & Farrell,H.M. : (1978) Interaction of ß-lactoglobulin and cytochrome c: complex formation and iron reduction. *Arch. Biochem. Biophys.* **185** 156-164

Brown,E.M., Pfeffer,P.E., Kumosinski,T.F. & Greenberg,R. : (1988) Accessibility and mobility of lysine residues in ß-lactoglobulin. *Biochem.* **27** 5601-5610

Bull,H.B. : (1946a) Monolayers of ß-lactoglobulin. II. Film molecular weight. *J. Am. Chem. Soc.* **68** 745-747

Bull,H.B. : (1946b) Monolayers of ß-lactoglobulin. II. Mixed monolayers of ß-lactoglobulin and sodium lauryl sulphate *J. Am. Chem. Soc.* **68** 747-748

Bull,H.B. & Currie,B.T. : (1946) Osmotic pressure of ß-lactoglobulin solutions. *J. Am. Chem. Soc.* **68** 742-745

Byler,D.M., Susi,H. & Farrell(Jnr),H.M. : (1983) Laser-raman spectra, sulphydryl groups, and conformation of the cystine linkages of ß-lactoglobulin. *Biopolymers* **22** 2507-2511

Cannan,R.K., Palmer,A.H. & Kibrick,A.C. : (1942) Hydrogen ion dissociatin curve of ß-lactoglobulin. *J. Biol. Chem.* **142** 803-822

Casal,H.L., Kohler,U. & Mantsch,H.N. : (1988) Structural and conformational changes of ß-lactoglobulin B: an infrared spectroscopic study of the effect of pH and temperature. *Biochim. Biophys. Acta* **957** 11-20

Cecil,R. & Ogston,A.G. : (1949) The sedimentation constant, diffusion constant and molecular weight of lactoglobulin. *Biochem J.* **44** 33-35

Christensen, L.K. : (1952) Denaturation and enzymic hydrolysis of lactoglobulin *C.R. Lab. Carlsberg (Ser Chim)* **28** 37-174

Clark,A.J., Simons,P., Wilmut,I. & Lathe,R. : (1987) Pharmaceuticals from transgenic

livestock. *TIBTech.* **5** 20-24
Conti,A. & Godovac-Zimmermann,J. : (1990) Antibodies raised against bovine ß-lactoglobulin react with ß2-microglobulin. *Biol. Chem. Hoppe-Seyler* **371** 261-263
Conti,A., Godovac-Zimmermann,J., Liberatori,J. & Braunitzer,G. : (1984) The primary structure of monomeric ß-lactoglobulin I from horse colostrum (*Equus caballus*, Perissodactyla). *Z. Physiol. Chem.* **366** 1393-1401
Conti,A., Godovac-Zimmermann,J., Pirchner,F., Liberatori,J. & Braunitzer,G. : (1986a) Pig ß-lactoglobulin I (*Sus scrofa domestica*, Artiodactyla). *Biol. Chem. Hoppe-Seyler* **367** 871-878
Conti,A., Liberatori,J. & Napolitano,L. : (1980) Isolation and preliminary physico-chemical characterization of human ß-lactoglobulin. *Milchwissenschaft* **35** 65-68
Conti,A., Liberatori,J. & Napolitano,L. : (1982a) Heterogeneity of equine ß-lactoglobulins. *Milchwissenschaft* **37** 338-340
Conti,A., Napolitano,L., Cantisani,A.M., Davoli,R. & Dall'Olio,S. : (1988) Bovine ß-lactoglobulin H : isolation by preparative isoelectric focusing in immobilized pH gradients and preliminary characterization. *J. Biochem. Biophys. Meth.* **16** 205-214
Conti,A., Napolitano,L., Godovac-Zimmermann,J., Liberatori,J. & Braunitzer,G. : (1986b) Preparative separation of pig polymorphic ß-lactoglobulins by IPG and non-IPG. *Protides Biol. Fluids Proc. Colloq.* **34** 907-910
Cooper,R., Eckley,D.M. & Papaconstantinou,J. : (1987) Nucleotide sequence of the mouse α_1-acid glycoprotein gene 1. *Biochemistry* **26** 5244-5250
Cornell, D.G. & Patterson, D.L. : (1989) Interaction of phospholipids in monolayers with ß-lactoglobulin adsorbed from solution. *J. Agric. Food Chem.* **37** 1455-1459
Creamer,L.K., Parry,D.A.D. & Malcolm,G.N. : (1983) Secondary structure of bovine ß-lactoglobulin B. *Arch. Biochem. Biophys.* **227** 98-105
Creighton,T.E. : (1980) Kinetic study of protein unfolding and refolding using urea gradient electrophoresis. *J. Mol. Biol.* **137** 61-80
Crowfoot,D. : (1941) A recent review of some recent X-ray work on protein crystals. *Chem. Rev.* **28** 215-228
Crowfoot,D.M. & Riley,D.P. : (1938) An x-ray study of Palmer's lactoglobulin. *Nature* **141** 521-522
Crowfoot,D.M. & Riley,D.P. : (1939) X-ray measurements on wet insulin crystals. *Nature* **144** 1011-1102
Cunningham,L.W. & Nuenke,B.J. : (1960) Analysis of modified ß-lactoglobulins and ovalbumins prepared from the sulfenyl iodide intermediates. *J. Biol. Chem.* **235** 1711-1715
Cupo,J.F. & Pace,C.N. : (1983) Conformational stability of mixed disulphide derivatives of ß-lactoglobulin B. *Biochemistry* **22** 2654-2658
Davies, D.T. : (1974) The quantitative partition of the albumin fraction of milk serum proteins by gel chromotography. *J. Dairy Res.* **41** 217-228
de Wit & Klarenbeck, G. : (1981) A differntial scanning calorimetric study of the thermal behaviour of bovine ß-lactoglobulin at temperatures up to 160°C. *J. Dairy Res* **48** 293-302
Deckmyn,H. & Preaux,G. : (1978) Chain folding prediction of the bovine ß lactoglobulins. *Arch. Int. Physiol. Biochim.* **86** 938-939
Diaz de Villegas,M.C., Oria,R., Sala,F.J. & Calvo,M. : (1987) Lipid binding by ß lactoglobulin of cow milk. *Milchwissenschaft* **42** 357-358
Drayna,D., Fielding,C., McLean,J., Baer,B., Casro,G., Chen,E., Comstock,L., Henzel,W., Kohr,W., Rhee,L., Wion,K. & Lawn,R. : (1986) Cloning and expression of human apolipoprotein D cDNA. *J. Biol. Chem.* **261** 16535-16539
Dunnill,P. & Green,D.W. : (1965) Sulphydryl groups and the N-R conformational change in ß-lactoglobulin. *J. Mol. Biol.* **15** 147-151
Dupont,M. : (1965) Etude d'une etape reversible dans la thermo- denaturation de la ß-lactoglobuline bovine A. *Biochim. Biophys. Acta* **102** 500-513
Erhardt,G., Godovac-Zimmerman,J. & Conti,A. : (1989) Isolation and complete primary sequence of a new ovine wild-type ß-lactoglobulin C. *Biol. Chem. Hoppe-Seyler* **370** 757-762
Fallstrom,S.P., Ahlstedt,S. & Hanson,L.A. : (1978) Specific antibodies in infants with gastrointestinal intolerance to cow's milk protein. *Int. Arch. All. Appl. Immun.* **56** 97-105
Farrell,H.M., Bede,M.J. & Enyeart,J.A. : (1987) Binding of p-nitrophenyl phosphate and other aromatics by ß-lactoglobulin. *J. Dairy Sci.* **70** 252-258
Farrell,H.M.(Jr) & Thompson,M.P. : (1971) Biological significance of milk protein polymorphism. *J. Dairy Sci.* **54** 1219-1228
Fernandez,F.M. & Oliver,G. : (1988) Proteins present in llama milk. I. Quantitative aspects and general characteristics. *Milchwissenschaft* **43** 299-302
Ferry, J.D. & Oncley, J.L. : (1941) Studies on the dielectric properties of protein solutions. III Lactoglobulin. *J. Amer. Chem. Soc.* **63** 272-278
Frank,G. & Braunitzer,G. : (1967) Der primarstruktur der ß-lactoglobuline. *Hoppe Seyler's Z. Physiol. Chem.* **348** 1691-1692
Freedman,R.B., Bullied, N.J., Hawkins,H.C. & Paver,J.L. : (1989) Role of protein disulphide-isomerase in the expresion of native proteins. *Biochem. Soc. Symp.* **55** 167-192
Frick,M. & Reiger,C.H.L. : (1987) Local antibodies to α-Casein and ß-lactoglobulin in saliva of infants. *Pediat. Res.* **22** 399-401
Frushour,B.G. & Koenig,L.J. : (1975) Raman studies of the crystalline, solution and alkaline-denatured states of ß-lactoglobulin. *Biopolymers* **14** 649-662
Fugate,R.D. & Song,P.-S. : (1980) Spectroscopic characterization of ß-lactoglobulin-retinol complex. *Biochim. Biophys. Acta* **625** 28-42
Futtermann,S. & Heller,J. : (1972) The enhancement of fluorescence and the decreased susceptibility to enzymatic oxidation of retinol complexed with bovine serum albumin, ß-lactoglobulin and the retinol-binding protein of human plasma. *J. Biol. Chem.* **247** 5168-5172
Garrick,J.D. & Williamson,I.W. : (1986) Is the role of ß-lactoglobulin to serve as a retinol carrier in milk? *BSc Honours Project Thesis*, University of Aberdeen
Gawinowicz,M.A. & Goodman,D.S. : (1985) The site of a retinoid affinity label to plasma retinol-binding protein. *J. Prot. Chem.* **4** 199-213
Gaye,P., Hue-Delahaie,D., Mercier,J.-C., Soulier,S., Vilotte,J.L. & Furet,J.P. : (1986) Ovine ß-lactoglobulin messenger RNA: nucleotide sequence and mRNA levels during functional differentiation of the mammary gland. *Biochimie* **68** 1097-1107
Georges,C., Guinand,S. & Tonnelat,J. : (1962) Etude thermodynamique de la dissociation reversible de la ß-lactoglobuline B pour des pH superieurs a 5.5. *Biochim. Biophys. Acta* **59** 737-739
Georges,C. & Guinand,S. : (1960) Sur la dissociation reversible de la $ß_2$-lactoglobuline, a des pH superieurs a 5.5 1. Etude par la diffusion de la lumiere. *J. Chem. Phys.* **57** 606-614
Ghose, A.C., Chaudhuri, S. & Sen, A. : (1968) Hydrogen ion equilibria and sedimentation behaviour of goat ß-lactoglobulin's. *Arch. Biochem. Biophys.* **126** 232-243
Godovac-Zimmermann,J. : (1988) The strucural motif of ß-lactoglobulin and retinol binding protein: a basic framework for binding and transport of small hydrophobic

molecules? *TIBS* **13** 64-66
Godovac-Zimmermann,J. & Braunitzer,G. : (1987) Modern aspects of the primary structure and function of ß-lactoglobulins. *Milchwissenschaft* **42** 294-297
Godovac-Zimmermann,J., Conti,A., James,L. & Napolitano,L. : (1988) Micro-analysis of the amino-acid sequence of monomeric ß-lactoglobulin I from donkey (*Equus asinus*) milk. *Biol. Chem. Hoppe-Seyler* **369** 171-179
Godovac-Zimmermann,J., Conti,A., Liberatori,J. & Braunitzer,G. : (1985a) Homology between the primary structures of ß-lactoglobulins and human retinol binding protein: evidence for a similar biological function? *Biol. Chem. Hoppe-Seyler* **366** 431-434
Godovac-Zimmermann,J., Conti,A., Liberatori,J. & Braunitzer,G. : (1985b) The amino-acid sequence of ß-lactoglobulin II from horse colostrum (*Equus caballus*, Perissodactyla) : ß-lactoglobulins are retinol-binding proteins. *Biol. Chem. Hoppe-Seyler* **366** 601-608
Godovac-Zimmermann,J., Conti,A. & Napolitano,L. : (1987) The complete amino-acid sequence of dimeric ß-lactoglobulin from mouflon (*Ovis ammon musimon*) milk. *Biol. Chem. Hoppe-Seyler* **368** 1313-1319
Godovac-Zimmermann,J., Krause,I. & Buchberger,J. : (1990) A novel wild-type ß-lactoglobulin W and its primary sequence *Biol. Chem. Hoppe-Seyler* **37** 255-260
Godovac-Zimmermann,J. & Shaw,D. : (1987) The primary structure, binding site and possible function of ß-lactoglobulin from Eastern grey kangaroo (*Macropus giganteus*). *Biol. Chem. Hoppe-Seyler.* **368** 879-886
Gorbunoff,M.J. : (1967) Exposure of tyrosine residues in proteins. Reaction of cyanuric fluoride with ribonuclease, α-lactalbumin, and ß-lactoglobulin. *Biochemistry* **6** 1606-1615
Granato,D.A. & Giguet,P.F. : (1986) A mouse monoclonal IgE antibody anti-bovine milk ß-lactoglobulin allows studies of allergy in gastrointestinal tract. *Clin. Exp. Immun.* **63** 703-710
Green,D.W. & Aschaffenburg,R. : (1959) Twofold symmetry of the ß-lactoglobulin molecule in crystals *J. Mol. Biol.* **1** 54-64
Green,D.W., Aschaffenburg,R., Camerman,A., Coppola,J.C., Diamand,R.D., Dunnill,P., Simmons,R.M., Komorowski,E.S., Sawyer,L., Turner,E.M.C. & Woods,K.F. : (1979) Structure of bovine ß-lactoglobulin at 6Å resolution. *J. Mol. Biol.* **131** 375-397
Green,D.W., North,A.C.T. & Aschaffenburg,R. : (1956) Crystallography of ß-lactoglobulin from cows milk. *Biochim. Biophys. Acta* **21** 583-585
Greenberg,R. & Kalan,E.B. : (1965) Studies on ß-lactoglobulins A,B and C. II. Preparation of modified proteins by treatment with carboxypeptidase A. *Biochemistry* **4** 1660-1666
Greene,R.F. & Pace,C.N. : (1974) Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, α-chymotrypsin and ß-lactoglobulin. *J. Biol. Chem.* **249** 5388-5393
Gronwall,A. : (1942) Solubility of ß-lactoglobulin in dilute NaCl solutions of various ionic strength and different degrees of pH. *C.R. Lab. Carlsberg (Ser. Chim.)* **24** 185-200
Grosclaude,F., Mahe,M.-F., Mercier,J.-C., Bonnemarie,J. & Teissier,J.H. : (1976) Polymorphisme des lactoproteines de bovines nepalais. *Ann. Genet. Sel. anim* **8** 461-479
Groves,M.L., Hipp,N.J. & McMeekin,T.L. : (1951) Effect of pH on the denaturation of ß-lactoglobulin and its dodecyl sulphate derivative. *J. Am. Chem. Soc.* **73** 2790-2793
Gutfreund, H. : (1945) The molecular weght of Palmers ß-lactoglobulin. *Nature* **155** 237 only.
Haefliger,J.-A., Jenne,D., Stanley,K.K. & Tschopp,J. : (1987) Structural homology of human complement component C8γ and plasma protein HC: identity of the cysteine bond pattern. *Biochem. Biophys. Res. Comm.* **149** 750-754
Halpin,M.I. & Richardson,T. : (1985) Selected functionality changes of ß-lactoglobulin upon esterification of side chain carboxyl groups. *J. Dairy. Sci.* **68** 3189-3198
Halwer,M., Nutting,G.C. & Brice,B.A. : (1951) Molecular weight of lactoglobulin, ovalbumin, lysozyme and serum albumin by light scattering. *J. Am. Chem. Soc.* **73** 2786-2790
Haque,Z., Kristjannson,M.M. & Kinsella,J.E. : (1987) Interaction between K-casein and ß-lactoglobulin: possible mechanism. *J. Agri. Food Chem.* **35** 644-649
Heller, W. & Klevens, H.B. : (1946) molecular weights of proteins as determined by light scattering. *Phys. Rev.* **67** 61 only
Hemley,R., Kohler,B.E. & Siviski,P. : (1979) Absorption spectra for the complexes formed from vitamin-A and ß-lactoglbulin. *Biophysics J.* 28 447-455
Hennighausen,L.G. & Sippel,A.E. : (1982) Mouse whey acidic protein is a novel member of the family of 'four-disulphide core' proteins. *Nucl. Acid Res.* **10** 2677-2684
Herskovits,T.T., Townend,R. & Timasheff,S.N. : (1964) Molecular interactions in ß-lactoglobulin. IX. Optical rotatory dispersion of the genetic variants in different states of association. *J. Am. Chem. Soc.* **86** 4445-4452
Hilquist,A.J. Dan,mo & Kresheck,G.C. : (1982) Influence of surfactants on the conformation of ß-lactoglobulin B using circular dichroism. *Biopolymers* **21** 895-908
Hin Bon Hoa, G., Douzou, P. & Pantaloni, C. : (1973) TRansformations alcalines de la ß-lactoglobulies en milieau hydro-alcoolique a basses temperatures. *Biochemie* **55** 269-276
Hodgkin,D.C. & Riley,D.P. (1968) Some ancient history of protein X-ray analysis *Structural Chemistry and Molecular Biology* Eds Rich,A. & Davidson,N. pp 15-28. Freeman, New York and London.
Holzwarth,G. : (1972) Ultraviolet spectroscopy of biological membranes. *Membrane Molecular Biology* **Chap 9** 228-286
Horisberger, M. & Vauthey, M. : (1984) Labelling of colloidal gold with protein: A quantitative study using ß-lactoglobulin. *Histochemistry* **22** 123
Horwitz,J. & Heller,J. : (1974) Properties of the chromophore binding site of retinol-binding protein from human plasma. *J. Biol. Chem.* **249** 4712-4719
Huang,Q., Coleman,J.W. & Stanworth,D.R. : (1985) Investigation of the allergenicity of ß-lactoglobulin and its cleavage fragments. *Int. Arch. All. Appl. Immun.* **78** 337-345
Huber,R., Schneider,M., Mayer,I., Muller,R., Deutzmann,R., Suter,F., Zuber,H., Falk,H. & Kayser,H. : (1987b) Molecular structure of the bilin binding protein (BBP) from *Pieris brassicae* after refinement at 2.0Å *J. Mol. Biol.* **198** 499-513
Hunziker,H.G. & Tarassuk,N.P. : (1965) Chromatographic evidenvce for heat induced interaction of α-lactalbumin and ß-lactoglobulin. *J. Dairy Sci.* **48** 733-744
Ivanov,V.N., Judinkova,E.S., Gorodetsky,S.I. : (1988) Molecular cloning of bovine ß-lactoglobulin cDNA. *Biol. Chem. Hoppe-Seyler* **369** 425-429
Jakobsson,I., Lindberg,T., Benediktsson,B. & Hansson,B.-G. : (1985) Dietary bovine ß-lactoglobulin is transferred to human milk. *Acta Ped. Scand.* **74** 342-345
Jamieson,A.C., Vandeyar,M.A., Kang,Y.C., Kinsella,J.E. & .A.Batt, C : (1987) Cloning and nucleotide sequence of the bovine ß-lactoglobulin gene. *Gene* **61** 85-90
Jenness,R. :(1974) The composition of milk. *Lactation* **Vol III chapter 1**, 3-107. Eds, Larson, B.L. & Smith, V.R., Academic Press NY/London.
Jenness,R. (1979) Comparative aspects of milk proteins *J.Dairy Res.* **46** 197-210
Jenness,R. :(1985) Biochemical and nutritional aspects of milk and colostrum. *Lactation* chapter 5 Ed. Larson, B.L. Iowa state UP 1985
Jenness,R., Phillip,N.I. & Kalan,E.B. : (1967) Immunochemical comparison of ß-lactoglobulins. *Fed. Proc.* **26** 340 only
Jones,M.N. & Wilkinson,A. : (1976) The interaction between ß-lactoglobulin and sodium

n-dodecyl sulphate.*Biochem. J.* **153** 713-718
Jones,T.A., Bergfors,T., Sedzik,J. & Unge,T. : (1988b) The three-dimensional structure of P2 myelin protein *Z.Krist.* **185** 120 only.
Julkunen,M., Seppala,M. & Janne,O.A. : (1988) Complete amino acid sequence of human placental protein 14: A progesterone-regulated uterine protein homologous to ß-lactoglobulins. *Proc. Natl. Acad. Sci.* **85** 8845-8849
Kalan,E.B. & Basch,J.J. : (1969) Preparation of goat ß-lactoglobulin. *J. Dairy Sci.* **49** 406-409
Kalan, E.B., Greenberg,R. & Walter,M. : (1965) ß-lactoglobulins A,B, and C. I. Comparison of chemical properties. *Biochem.* **4** 991-997
Kalan,E.B., Kraeling,R.R. & Gerrits,R.J. : (1971) Isolation and partial characterization of a polymorphic swine whey protein. *Int. J. Biochem.* **2** 232-244
Kamur,M., Dastur,N.N. & Bhatt,M.V. : (1987) Studies on the pigments of buffalo milk. I.Identification of biliverdin IXa isomer in fresh buffalo milk. *Milchwissenschaft* **39** 29-31
Keen, J.N., Caceres, I., Eliopolous, E.E., Zagalsky, P.F. & Findlay, J.B.C. : (1990) Complete sequence and model for the A_2-subunit of the carotenoid pigment complex, crusticyanin. *Eur. J. Biochem.* In Press.
Kella, & Kinsella, (1988) Enhanced thermodynamic stabilityof ß-lactoglobulin at low pH. *Biochem. J.* **255** 113-118
Kessler,E. & Brew,K. : (1970) The whey proteins of pig's milk isolation and characterization of a ß-lactoglobulin. *Biochim. Biophys. Acta* **200** 449-458
Kiddy, C.A., Rollins, R.E. & Zikakis, J.P. : (1972) Discontinuous polyacrylamide electrophoresis for ß-lactoglobulin of cow milk. *J. Dairy Sci.* **55** 1506-1507
Kolde,H.-J. & Braunitzer,G. : (1983) The primary structure of ovine ß-lactoglobulin. 1. Isolation of the peptides and sequence. *Milchwissenschaft* **38** 18-20
Koritz,T.N., Suzuki,S. & Coomb,R.R.A. : (1987) Antigenic stimulation with proteins of cows milk via the oral route in guinea-pigs and rats. 1. Measurement of antigenically intact ß-lactoglobulin and casein in the gastrointestinal contents of duodenum, jejunum and ileum. *Int. Arch. All. Appl. Immun.* **82** 72-75
Kraeling,R.R. & Gerrits,R.J. : (1969) Polymorphism of a protein of sows' whey. *J. Dairy Sci.* **52** 2036-2038
Kurisaki,J., Nakamura,S., Kaminogawa,S. & Yamauchi,K. : (1982) The antigenic properties of ß-lactoglobulin examined with mouse IgE antibody. *Agri. Biol. Chem.* **46** 2069-2075
Kurisaki,J., Nakamura,S., Kaminogawa,S., Yamauchi,K., Watanabe,S., Hotta,K. & Hattori,M. : (1985) Antigenicity of modified ß-lactoglobulin examined by three different assays. *Agri. Biol. Chem.* **49** 1733-1737
Larson,B.L. : (1979) Biosynthesis and secretion of milk proteins: a review. *J.Dairy Sci.* **46** 161-174
Larson,B.L. & Jenness,R. : (1952) Characterization of the sulphydryl groups and the kinetics of the heat denaturation of crystalline ß-lactoglobulin. *J. Am. Chem. Soc.* **74** 3090-3093
Lebenthal,E., Laor,J., Lewitus,Z., Matoth,Y. & Freier,S. : (1970) Gastrointestinal protein loss in allergy to cows milk ß-lactoglobulin. *Isr. J. Med. Sci.* **6** 506-510
Lee,K.-H., Wells,R.G. & Read,R.R. : (1987) Isolation of an olfactory cDNA: similarity to retinol-binding protein suggests a role in olfaction. *Science* **235** 1053-1056
Li,C.H. : (1946) Electrophoretic inhomogeneoty of crystalline ß-lactoglobulin. *J. Amer. Chem Soc.* **68** 2746-2747
Liberatori,J., Guidetti,L.M. & Conti,A. : (1979a) Immunochemical studies on ß-lactoglobulins. Precipitin reactions of sow's and mare's mammary secretions against anti-bovine ß-lactoglobulin antiserum. *Boll. Soc. It. Biol. Sper.* **55** 815-821
Liberatori,J., Morisio,L. Guetti,id, Conti,A. & Napolitano,L. : (1979b) ß-lactoglobulins in the mammary secretions of camel (*Camelus dromedarius*) and she-ass. Immunological detection and preliminary physico-chemical characterization. *Boll. Soc. It. Biol. Sper.* **55** 1369-1373
Lockwood,D.H., Turkington,R.W. & Topper,Y.J. : (1966) Hormone- dependent development of milk protein synthesis in mammary gland in vitro. *Biochim. Biophys. Acta* **130** 493-501
Lontie,R. & Preaux,G. : (1966) Polarimetric investigation of ß-lactoglobulin A and B and of the reactivity of their thiol groups. *Protides Biol. Fluids 14* 475-482
Lovrien,R. & Anderson,W.F. : (1969) Resolution of binding sites in ß-lactoglobulin. *Arch. Biochem. Biophys.* **131** 139-144
Lyster,R.L.J. : (1964) The reactivity of the -SH groups of ß-lactoglobulin subunits. *Abst. 6th Int. Congress Biochem.*
Lyster,R.L.J. : (1972) Reviews of the progress of dairy science. Section C. Chemistry of milk proteins. *J. Dairy Res.* **39** 279-318
Lyster,R.L.J., Jenness,R., Phillips,N.I. & Sloan,R.E. : (1966) Comparative biochemical studies of milks - III. Immuno- electrophoretic comparisons of milk proteins of the Artiodactyla. *Comp. Biochem. Physiol.* **17** 967-971
Lyster,R.L.J. & Wheelock,J.V. : (1967) Occurrence of milk proteins in urine of cows during an extended milking interval. *J. Dairy Res.* **34** 27-30
Macheboeuf,M. & Robert,B. : (1953) Etude sur la denaturation des proteines 1. Autoxydation des proteines en milieu alcalin. *Bull. Soc. Chim. Biol.* **35** 399-411
Maillert, P. & Ribadeau-Dumas, B. : (1988) Preparation of ß-lactoglobulin and ß-lactoglobulin-free proteins from whey retentate by NaCl salting out at low pH *J. Food Sci.* **53** 743-745
Mainferme,F., Preaux,G. & Lontie,R. : (1971) Location of the disulphide bridges in the sequence of ß-lactoglobulin A with CNBr and thermolysin. *Arch. Int. Physiol. Biochem.* **79** 840-841
Malik,Z., Bottomley,R. & Austen,B. : (1988) Allergenic properties of the genetic variants A and B of bovine ß-lactoglobulin. *Int. Arch. All. Appl. Immun.* **86** 245-248
Martial,J., Preaux,G. & Lontie,R. : (1971) Location of the thiol group and of the disulphide bridges in the sequence of ß-lactoglobulin B. *Arch. Int. Physiol. Biochem.* **79** 842-843
Mattarella,N.L. & Richardson,T.L. : (1983) Physicochemical and functional properties of positively charged derivatives of bovine ß-lactoglobulin. *J. Agric. Food Chem.* **31** 972-978
Maubois,J.L., Pion,R. & Ribadeau-Dumas,B. : (1965) Preparation et etude de la ß-lactoglobulin de brebis crystallisee. *Blochim. Biophys. Acta* **107** 501-510
McAlpine,A.S. & Sawyer,L. : (1990) ß-Lactoglobulin: a protein drug carrier? *Biochem.Soc.Trans.* In press
McDougall,E.I. & Stewart,J.C. : (1976) The whey proteins of the milk of red deer (*Cervus elaphus L.*). *Biochem. J.* **153** 647-655
McDougall,E.I. & Stewart,J.C. : (1977) The state of aggregation of red deer (*Cervus elaphus L.*) ß-lactoglobulin preparations near neutral pH. *Biochem. J.* **167** 45-51
McKenzie,H.A. : (1967) Milk Proteins. *Adv. in Prot. Chem.* **22** 147-234
McKenzie,H.A. : (1971) ß-lactoglobulins. *Milk Proteins II* 257-330
McKenzie,H.A., Muller,V.J. & Treacy,G.B. : (1983) "Whey" proteins of milk of the red (*Macropus rufus*) and Eastern grey (*Macropus giganteus*) kangaroo. *Comp. Biochem. Physiol.* **74B** 259-271
McKenzie,H.A. & Ralston,G.B. : (1971) The denaturation of proteins : two states? reversible or irreversible? *Experientia* **27** 617-624

McKenzie,H.A. & Ralston,G.B. : (1973) Nature of products formed by the action of urea on bovine ß-lactoglobulins. *Aust. J. Biol. Sci.* **26** 859-870
McKenzie,H.A., Ralston,G.B. & Shaw,D.C. : (1972) Location of sulphydryl and disulphide groups in bovine ß-lactoglobulins and effects of urea. *Biochemistry* **11** 4539-4547
McKenzie,H.A. & Sawyer,W.H. : (1967) Effect of pH on ß-lactoglobulins. *Nature* **214** 1101-1104
McKenzie,H.A., Sawyer,W.H. & Smith,M.B. : (1967) Optical rotatory dispersion and sedimentation in the study of association- dissociation : bovine ß-lactoglobulins near pH 5. *Biochim. Biophys. Acta* **147** 73-92
McKenzie,H.A. & Shaw,D.C. : (1972) Alternative positions for the sulphydryl group in ß-lactoglobulin: the significance for sulphydryl location in other proteins. *Nature* **238** 147-149
McMeekin,T.L., Polis,B.D., Della-Monica,E.S. & Custer,J.H. : (1949) A crystallization complex of ß-lactoglobulin with dodecyl sulphate. *J. Am. Chem. Soc.* **71** 3606-3609
McMeekin,T.L. & Warner,R.C. : (1942) The hydration of ß-lactoglobulin crystals. *J. Am. Chem. Soc.* **64** 2393-2398
McRee, D.E., Tainer, J.A., Meyer, T.E., Van Beeuman, J., Cusanovich, M.A. & Getzoff, E.D. : (1989)Crystallographic structure of a photoreceptor protein at 2.4 Å resolution. *Proc. Nat. Acad. Sci.* **86** 6533-6537
Mercier,J.-C. & Gaye,P. : (1982) Early events in secretion of main milk proteins: occurrence of precursors. *J.Dairy Sci.* **65** 299-316
Mercier,J.C., Gaye,P., Soulier,S., Hue-Delahaie,D. & Vilotte,J.L. : (1985) Construction and identification of recombinant plasmids carrying cDNAs coding for ovine αs1-, αs2-, ß-, κ-casein and ß-lactoglobulin. Nucleotide sequence of αs1-casaein cDNA. *Biochemie* **67** 959-971
Mills,O.E. : (1976) Effect of temperature on tryptophan fluorescence of ß-lactoglobulin B. *Biochim. Biophys. Acta* **434** 324-332
Mills,O.E. & Creamer,L.K. : (1975) Conformational changes of bovine ß-lactoglobulin at low pH. *Biochem. Biophys. Acta* **379** 618-626
Miranda,G. & Pelissier,J.-P. : (1983) Kinetic studies of in vivo digestion of bovine unheated skim-milk proteins in rat stomach *J.Dairy Res.* **50** 27-36
Mohammadzadeh-K.,A., Feeney,R.E. & Smith,L.M. : (1969) Hydrophobic binding of hydrocarbons by proteins. I. Relationship of hydrocarbon structure. *Biochim. Biophys. Acta* **194** 246-255
Monaco,H.L., Zanotti,G., Spadon,P., Bolognesi,M., Sawyer,L. & .E.Eliopoulos, E : (1987) Crystal structure of the trigonal form of bovine ß-lactoglobulin and its complex with retinol at 2.5 Å resolution. *J. Mol. Biol.* **197** 695-706
Monti, J.C., Mermoud, A.-F. & Jolles, P. : (1989) Anti-bovine ß-lactoglobulin antibodies react with a human lactoferrin fragment and bovine ß-lactoglobulin present in human milk. *Experentia* **45** 178-180
Newcomer,M.E., Jones,T.A., Aqvist,J., Sundelin,J., Erikkson,U., Rask,L. &P.Peterson,.A : (1984) The three dimensional structure of retinol-binding protein. *EMBO J.* **3** 1451-1454
North,A.C.T. : (1989a) Three-dimensional arrangement of conserved amino acid rersidues in a superfamily of specific ligand-binding proteins. *Int. J. Biol. Macromol.* **11** 56-58
North,A.C.T. : (1989b) Application of molecular graphics for the study of recognition. *J. Mol. Graphics* **7** 67-70
Novotny,J. & Franek,F. : (1975) Different degrees of interspecies homology in immunoglobulin lambda chain constant domain correlated with three-dimensional structure. *Nature* **258** 641-643
O'Neill,T.E. & Kinsella,J.E. : (1987) Binding of alkanone flavors to ß-lactoglobulin: effects of conformational and chemical modification. *J. Agri. Food Chem.* **35** 770-774
Osterhoff,D.R. & Pretorius,A.M.G. : (1966) Inhereited biochemical polymorphism. *Proc. S. Afri. Soc. Anim. Prod.* **5** 166-173
Otani,H. : (1981) Antigenicities of ß-lactoglobulin treated with proteolytic enzymes. *Jap. J. Zootech. Sci.* **52** 47-52
Otani,H. & Hosono,A. : (1987) Antigenic reactive regions of S-carboxymethylated ß-lactoglobulin. *Agri. Biol. CHem.* **51** 531-536
Otani,H., Morita,S. & Tokita,F. : (1985a) Studies on the antigenicity of the browning product between ß-lactoglobulin and lactose. *Jap J. Zootech. Sci* **56** 67-74
Otani,H. & Tokita,F. : (1982) Studies on the antigenic structure of the browning product between ß-lactoglobulin and lactose by chemical fragmentation. *Jap. J. Zootech. Sci.* **53** 293-301
Otani,H., Uchio,T. & Tokita,F. : (1985b) Antigenic reactivities of chemically modified ß-lactoglobulins with antiserum to bovine ß-lactoglobulin. *Agri. Biol. CHem.* **49** 2531-2536
Palmer,A.H. : (1934) The preparation of a crystalline globulin from the albumin fraction of cow's milk. *J. Biol. Chem.* **104** 359-372
Pantaloni,D. : (1962) Role des groupes SH dans la structure de la ß-lactoglobuline. *Compte. Rend.* **254** 1884-1886
Pantaloni,D. : (1964) Etude de la transition R-S de la ß-lactoglobuline par spectropolarimetrie et par spectrophotometrie de differences. *Compte. Rend. Seanc. Acad. Sci. Paris* **258** 5753-5756
Pantaloni,D. : (1965) Structure et changements de conformations de la ß-lactoglobuline en solution. Thesis for doctorate : University of Paris
Papiz,M.Z. : (1982) Crystallographic studies of ß-lactoglobulin at 2.8Å resolution. Doctoral Thesis, Napier College, Edinburgh
Papiz,M.Z., Sawyer,L., Eliopoulos,E.E., North,A.C.T., Findlay,J.B.C., Sivaprasadarao,R., Jones,T.A., Newcomer,M.E. & Kraulis,P.J. : (1986) The structure of ß-lactoglobulin and its similarity to plasma retinol-binding protein. *Nature* **324** 383-385
Park,K.H. & Lund,D.B. : (1984) Calorimetric study of the thermal denaturation of ß-lactoglobulin. *J. Dairy. Sci.* **67** 1699-1706
Pederson,K.O. : (1936a) Introduction and preliminary results with fractionation of skim milk. *Biochem,. J.* **30** 948-960
Pederson,K.O. : (1936b) The lactoglobulin of Palmer: Detected changes sedimentation coefficient at pH 5.0 and 7.5. *Biochem,. J.* **30** 961-970
Peitsch, M.C. & Boguski, M.S. : (1990) Is apolipoprotein D a mammalian bilin-binding protein? *The New Biologist* **2** 197-206
Perez,M,D., Diaz de Villegas,C., Sanchez,L., Arando,P., Ena,J.M. & Calvo,M. : (1989) Interaction of fatty acids with ß-lactoglobulin and albumin from ruminant milk *J.Biochem.* **106** 1094-1097
Perez-Gomez,G., Preaux,G. & Lontie,R. : (1971) Location of the disulphide bridges in ß-lactoglobulin B with pepsin and trypsin after CNBr cleavage. *Arch. Int. Physiol. Biochem.* **79** 843-844
Pervaiz,S. & Brew,K. : (1985) Homology of ß-lactoglobulin, serum retinol-binding protein, and protein HC. *Science* **228** 335-337
Pervaiz,S. & Brew,K. : (1986) Purification and characterization of the major whey proteins from the milks of the Bottlenose dolphin (*Tursiops truncatus*), the Florida manatee (*Trichechus manatus latirostris*) and the beagle (*Canis familiaris*). *Arch. Biochem.*

Biophys. **246** 846-854
Pervaiz,S. & Brew,K. : (1987) Homology and structure-function correlations between α_1-acid glycoprotein and serum retinol-binding protein and its relatives *FASEB* **1** 209-214
Pierce, A.E. : (1960) ß-Lactoglobulins in the urine of the new-born suckled calf. *Nature* **186** 940-941
Piez,K.A., Davie,E.W., Folk,J.E. & Gladner,J.A. : (1961) ß-lactoglobulins A and B. *J. Biol. Chem.* **236** 2912-2916
Polis,P.D., Schmuckler,H.W., Custer,J.H. & McMeekin,T.L. : (1950) Isolation of an electrophoretically homogenous crystalline component of ß-lactoglobulin *J. Amer. Chem. Soc.* **72** 4965-4968
Preaux,G., Braunitzer,G. & Kolde,H.-J. : (1980) Primary structure of ovine ß-lactoglobulin. *Arch. Int. Physiol. Biochem.* **88** B45 35-36
Preaux,G., Braunitzer,G., Schrank,B. & Stangl,A. : (1979) The amino acid sequence of goat ß-lactoglobulin. *Hoppe Seyler's Z. Physiol. Chem.* **360** 1595-1604
Preaux,G. & Lontie, R. : (1972) Revised number of the free cysteine residues and of the disulphide bridges in the sequence of ß-lactoglobulins A and B. *Arch. Int. Physiol. Biochim.* **80** 980-981
Prutz,W.A., Butler,J., Land,E.J. & Swallow,A.J. : (1980) Direct demonstration of electron transfer between tryptophan and tyrosine in proteins. *Biochem. Biophys. Res. Comm.* **96** 408-414
Purcell,J.M. & Susi,H. : (1984) Solvent denaturation of proteins as observed by resolution-enhanced Fourier transform infrared spectroscopy. *J. Biochem. Biophys. Meth.* **9** 193-199
Rask,L., Anundi,H. & Peterson,P.A. : (1979) The primary structure of human retinol binding protein. *FEBS. Lett.* **104** 55-58
Reithel, F.J.. & Kelly, M.J : (1971)Thermodynamic analysis of the monomer-dimer association of ß-lactoglbulin A at the isoelectric point. *Biochem* **10** 2639-2644
Riley,C.T., Barbeau,B.K., Keim,P.S., Kezdy,F.J., Heinrikson,R.L. & Law,J.H. : (1984) The covalent protein structure of insecticyanin, a blue biliprotein from the hemolymph of the tobacco hornworm, *Manduca sexta* L.. *J. Biol. Chem.* **259** 13159-13165
Robertson,D.M., Paganelli,R., Dinwiddie,R. & Levinsky,R.J. : (1982) Milk antigen absorption in the preterm and term neonates. *Arch. Dis. Child.* **57** 369-372
Robillard,K.A. : (1971) Hydrophobic interactions between small molecules and ß-lactoglobulin *A. Fed. Proc.* **30** 1269 only
Robillard,K.A. & Wishnia,A. : (1972b) Aromatic hydrophobes and ß-lactoglobulin A. Kinetics of binding by nuclear magnetic resonance. *Biochemistry* **11** 3841-3845
Roels,H., Preaux,G. & Lontie,R. : (1966) Stabilization of ß-lactogloblin A and B at pH 8.9 by blocking the thiol groups. *Arch. Int. Biochem.* **74** 522-523
Ruegg, M., Moor, U. & Blanc, B. : (1975) Hydration and thermal denaturation of ß-lactoglobulin. A calorimetric study. *Biochem. Biophys. Acta.* **400** 334-342
Sacchettini,J.C., Gordon,J.I. & Banaszak,L.J. : (1989) The structure of crystalline *Escherichia coli*- derived rat intestinal fatty acid-binding protein at 2.5Å resolution. *J. Biol. Chem.* **263** 5815-5819
Said,H.M., Ong,D.E. & Shingleton,J.L. : (1989) Intestinal uptake of retinol: enhancement by bovine milk ß-lactoglobulin *Am.J.Clin.Nutr.* **49** 690-694
Sawyer,L. : (1987) One fold among many. *Nature* **327** 659 only
Sawyer,L. & Green,D.W. : (1979) The reaction of cow ß-lactoglobulin with tetracyanoaurate(III). *Biochim. Biophys. Acta* **579** 234-239
Sawyer,L., Papiz,M.Z., North,A.C.T. & Eliopoulos,E.E. : (1985) Structure and function of bovine ß-lactoglobulin. *Biochem. Soc. Trans.* **13** 265-266
Sawyer,W.H. : (1969) Complex between ß-lactoglobulin and κ-casein. A review. *J. Dairy Sci* **52** 1347-1355
Schellman, J.A. : (1988) Rotatory properties of ß-lactoglobulin. *Compt. Rend. Trav. Lab. Carlsberg Ser. Chim.* **30** 395-414
Seibles,T.S. : (1969) Interaction of dodecyl sulphate with native and modified ß-lactoglobulin. *Biochemistry* **8** 2949-2954
Senti,F.R. & Warner,R.C. : (1948) X-ray molecular weight of ß-lactoglobulin *J.Amer.Chem.Soc.* **70** 3318-3320
Sen,A. & Chaudhuri,S. : (1962) Non-casein proteins of goat's milk. *Nature* **195** 286-287
Sen,A. & Sinha,N.K. : (1961) Comparison of the ß-lactoglobulins of buffalo's and cow's milk. *Nature* **190** 343-344
Simons,J.P., McClenaghan,M. & Clark,A.J. : (1987) Alteration of the quality of milk by expression of sheep ß-lactoglobulin in transgenic mice. *Nature* **328** 530-532
Smith,L.M., Fantozzi,P. & Creveling,R.K. : (1983) Study of triglyceride-protein interaction using a microemulsion-filtration method. *J. Am. Oil Chem. Soc.* **60** 960-967
Spector,A.A. & Fletcher,J.E. : (1970) Binding of long chain fatty acids to ß-lactoglobulin. *Lipids* **5** 403-411
Stone,W.L. & Wishnia,A. : (1978) Binding of iodomercurates to sulphydryl-blocked ß-lactoglobulin -A, -B, and -C. *Bioinorg. Chem.* **8** 517-529
Su,Y.T. & Jirgensons,B. : (1977) Further studies on detergent-induced conformational transitions in proteins. Circular dichroism of ovalbumin, bacterial α-amylase, papain, and ß-lactoglobulin at various pH values. *Arch. Biochem. Biophys.* **181** 137-146
Sutur,F., Kayser,H. & Zuber,H. : (1988) The complete amino-acid sequence of the bilin-binding protein from *Pieris brassicae* and its similarity to a family of serum transport proteins like retinol-binding proteins. *Biol. Chem. Hoppe-Seyler* **369** 497-505
Suzuki,S., Koritz,T.N. & Coomb,R.R. : (1987) Antigenic stimulation with proteins of cows milk via the oral route in guinea-pigs and rats.2. Antibodies to ß-lactoglobulin secreted into the alimentary canaland serum. *Int. Arch. All. Appl. Immun.* **82** 76-82
Svedberg (1938) The ultracentrifuge and its field of research. *Ind. Eng. Chem.* **10** 113-129
Swaisgood, H.E. : (1982) Chemistry of milk protein. *Developments in dairy chemistry* **1**, Ed Fox, P.F.
Tanford,C., Bunville,L.G. & Nozaki,Y. : (1959) The reversible trans-formation of ß-lactoglobulin at pH 7.5. *J. Am. Chem. Soc.* **81** 4032-4036
Tanford,C., De,P.K. & Taggart,V.G. : (1960) The role of the α-helix in the structure of proteins. Optical rotatory dispersion of ß-lactoglobulin. *J. Am. Chem. Soc.* **82** 6028-6034
Tanford,C. & Taggart,V.G. : (1961) Ionization-linked changes in protein conformation. II. The N-R transition in ß-lactoglobulin. *J. Am. Chem. Soc.* **83** 1634-1638
Teller, D.C., Swanson, E., DeHaen, C. : (1979) The translational frictional coefficient of proteins. *Mets. In Enzymol.* **61** 103-124
Thompson, M.P. : (1970) Phenotyping milk proteins: A review. *J. Dairy Sci.* **53** 1341-1348
Thompson,M.P. & Farrell,H.M. (Jr) : (1974) Genetic variants of milk proteins. Lactation. *A Comprhensive Treatise* **Vol III** 109-134
Tilley, J.M.A. : (1960) The chemical and physical properties of bovine ß-lactoglobulin. *Dairy Sci. Abs.* **22** 111-125
Timasheff,S.N., Mescanti,L., Basch,J.J. & Townend,R. : (1966b) Conformational transitions of bovine ß-lactoglobulins A, B, and C. *J. Biol. Chem.* **241** 2496-2501

Timasheff,S.N. & Townend,R. : (1961) Molecular interactions in ß-lactoglobulin. V. The association of the genetic species of ß-lactoglobulin below the isoelectric point. *J. Am. Chem. Soc.* **83** 464-469

Timasheff,S.N. & Townend,R. : (1964) Structure of the ß-lactoglobulin tetramer. *Nature* **203** 517-519

Timasheff,S.N., Townend,R. & Mescanti,L. : (1966a) The optical rotatory dispersion of ß-lactoglobulins. *J. Biol. Chem.* **241** 1863-1870

Timasheff,S.N. & Susi,H. : (1966) Infrared investigation of the secondary structure of ß-lactoglobulins. *J. Biol. Chem.* **241** 249-251

Townend,R., Herskovits,T.T., Timasheff,S.N. & Gorbunoff,M.J. : (1969) The state of amino acid residues in ß-lactoglobulin. *Arch. Biochem. Biophys.* **129** 567-580

Townend,R., Herskovits,T.T., Swaisgood, H.E. & Timasheff, S.N. : (1964) Solution properties of ß-lactoglobulin C. *J. Biol. Chem.* **239** 4196-4201

Townend,R., Kumosinski,T.F. & Timasheff,S.N. : (1967) The circular dichroism of variants of ß-lactoglobulin. *J. Biol. Chem.* **242** 4538-4545

Townend,R., Weinberger,L. & Timasheff,S.N. : (1960b) Molecular interactions in ß-lactoglobulin. IV. The dissociation of ß-lactoglobulin below pH 3.5. *J. Am. Chem. Soc.* **82** 3175-3179

Townend,R., Winterbottom,R.J. & Timasheff,S.N. : (1960a) Molecular interactions in ß-lactoglobulin. II. Ultracentrifugal and electrophoretic studies of the association of ß-lactoglobulin below its isoelectric point. *J. Am. Chem. Soc.* **82** 3161-3168

Townend,R. & Timasheff,S.N. : (1960) Molecular interactions in ß-lactoglobulin III. Light scattering investigation of the stoichiometry of the association between pH 3.7 and 5.2. *J. Am. Chem. Soc.* **82** 3168-3174

Treece, J.M., Sheinson, R.S. & McMeekin, T.L. : (1964) The solubilities of ß-lactoglobulins A,B, and AB. *Arch. Biochem. Biophys.* **108** 99-108

Unterman,R.D., Lynch,K.R., Nakhasi,H.L., Dolan,K.P., Hamilton,J.W., Cohn,D.V. & Feigelson,P. : (1981) Cloning and sequence of several α_2u-globulin cDNAs. *Proc. Nat. Acad. Sci.* **78** 3478-3482

Whitney, R.M. : (1974) Nomenclature of the proteins of cows milk: fourth revision. *J. Dairy Sci.* **59** 798-815

Willis,I.M., Stewart,A.F., Caputo,A., Thompson,A.R. & Mackinlay,A.G. : (1982) Construction and identification by partial nucleotide sequence analysis of bovine casein and ß-lactoglobulin cDNA clones. *DNA* **1** 375-386

Wishnia,A. & Pinder,T.W. (Jr) : (1966) Hydrophobic interactions in proteins. The alkane binding site of ß-lactoglobulins A and B. *Biochemistry* **5** 1534-1542

Woo,S.L., Creamer,L.K. & Richardson,T. : (1982) Chemical phosphorylation of bovine ß-lactoglobulin. *J. Agri. Food Chem.* **30** 65-70

Workman,R.J., McKown,M.M. & Gregerman,R.I. : (1974) Renin: inhibition by proteins and peptides. *Biochemistry* **13** 3029-3035

Yewdall,S. (1988) Structural studies on ß-lactoglobulin. PhD Thesis, University of Leeds.

Yoshikawa,M., Mizukami,T., Sasaki,R. & Chiba,H. : (1978) Pre-ß-lactoglobulin synthesis by mRNA from bovine mammary gland. *Agri. Biol. Chem.* **42** 2185-2186

Zimmerman, J.K., Barlow,G.H. & Klotz,I.M. : (1970) Dissociation of ß-lactoglobulin near neutral pH *Arch.Biochem.Biophys.* **138** 101-109

TABLE I

DISTRIBUTION OF BLG IN THE MILKS OF VARIOUS SPECIES.

Species	Cross-reaction with antibovine antisera	State of association
Cow (Bos taurus)	+	Di[a]
Oxen (Bos javanicus)	+	Di[b]
Yak (Bos grunniens)	+	Di
Zebu (Bos indicus)	+[c]	?Di*
Waterbuffalo (Bubalus arnee)	+	Di
Buffalo (Bubalis bubalis)	+[c]	?Di*
Bison (Bison bison)	+[c]	?Di*
Musk ox (Ovibos moschatus)	+[c]	?Di*
Eland (Taurotragus oryx)	+[c]	?Di*
Goat (Capra hircus)	+	Di
Sheep (Ovis aries)	+	Di[d]
Mouflon (Ovis ammon musimon)	+	Di[e]
Red deer (Cervus elaphus L.)	+[c]	?Di*
European elk (Alces alces L.)	+[c]	?Di*
Reindeer (Rangifer tarandus L.)	+[c]	?Di*
White-tailed deer (Odocoileus virginianus)	+[c]	?Di*
Fallow deer (Dama dama)	+[c]	?Di*
Caribou (Rangifer arcticus)	+[c]	?Di*
Giraffe (Giraffa camelopardalis)	+[c]	?Di*
Okapi (Okapia johnstoni)	+[c]	?Di*
Prong-horn antelope (Antilocapra americana)	+[c]	?Di*
Camel (Camelus dromedarius)	+/-[f]	?Mono*
Llama (Lama glama L.)	-[g]	none[g]
Peccary (Pecari tajacu)	-*[c]	Mono
Pig (Sus scrofa domestica)	+/-[h]	Mono
Horse (Equus caballus)	+/-[h, i]	Mono
Zebra (Equus quagga)	-[c]	?none/Mono*
Rhinoceros (Diceros bicornis)	-[c]	?none/Mono*
Donkey (Equinus asinus)	-*[j]	Mono[j]
Mouse (Mus musculus)	-	none[k]
Rat (Ratus norvegicus)	nm	none[l]
Guinea-pig (Cavia porcellus)	nm	none[m]
Dolphin (Tursiops truncatus)	nm	Mono[n]
Manatee (Trichechus manatus latirostris)	nm	Mono[n]
Beagle (Canis familiaris)	nm	Mono[n]
Cat (Felis catus)	-[o]	?Mono*
Grey kangaroo (Macropus giganteus)	nm	Mono*[p]
Red kangaroo (Macropus rufus)	nm	Mono[p]
Northern fur seal (Callorhinus ursinus)	nm	unknown[q]
Human (Homo sapiens)	+/-[r]	none*[s]
~~Chimpanzee (Pan troglodytes)~~	~~nm~~	~~none[19]~~
~~Lemur (Galago crassicaudatus)~~	~~nm~~	~~none[19]~~
Rabbit (Oryctolagus cuniculus)	-[c]	none

Key
+ = BLG detected by antibovine BLG antisera
- = BLG not detected by antibovine BLG antisera
+/- = BLG detected by antibovine BLG antisera, but only at higher titres
nm = cross-reactivity to antibovine BLG antisera not measured
Di = Dimeric BLG detected
Mono = Monomeric BLG detected
None = no BLG present
unknown = BLG present, but its state of association unknown
* = ~~one host~~, cross reactivity to antibovine BLG antisera is assumed to indicate the presence of a dimeric form of the protein

a. 87 - Bull et al., (1946).

b. 75 - Bell et al., (1981).

c. 17 - Lyster et al., (1966).

d. 77 - Bell & McKenzie, (1967).

e. 37 - Godovac-Zimmerman et al., (1987).

f. 18 - Libaratori et al., (1979).

g. 48 - Fernandez & Oliver, (1988).

h. 83b - Liberatori et al., (1979).

i. 42 - Bell et al., (1981).

j. 43 - Godovac-Zimmerman et al., (1988).

k. 29 - Clark et al., (1987).

l. 26 - Hennighausen & Sippel, (1982).

m. 47 - Brew & Campbell, (1967).

n. 44 - Pervaiz & Brew, (1986).

o. 31 - Bell, (1990), personal communication.

p. 14 - McKenzie et al., (1983).

q. 46 - Ashworth et al., (1966).

r. 20a - Conti et al., (1979).

s. 19 - Bell & McKenzie, (1964).

TABLE III

ISOLATION PROCEDURES FOR PIG BLGS.

Method	Removal of fat	Removal of caseins	Fractionation of whey proteins	Purification
Kessler & Brew (1970)	Acidification		Gel filtration	Ion-exchange chromatography Gel filtration
Kalan, Kraeling & Gerrits (1971)	Separator at 40°C	By acid precipitation	By sodium sulphate precipitation to remove aLa, acid precipitation of BLG at pH 3	pH fractionation Ammonium sulphate fractionation Crystallization
Bell, Shaw & McKenzie (1981)	Centrifugation	By ammonium sulphate precipitation	Ammonium sulphate precipitation of whey proteins, gel filtration	Ion-exchange chromatography

Key
aLa = α-lactalbumin

TABLE II

Some ISOLATION PROCEDURES FOR BOVINE BLGS.

Method	Removal of fat	Removal of caseins	Fractionation of whey proteins	Purification
Palmer (1954)	Skim milk used	Acid precipitation	By sodium sulphate precipitation at up to 30°C	Dialysis Recrystallization
Aschaffenburg & Drewry (1957)	With sodium sulphate precipitation at 40°C		Acid precipitation at pH 2 to remove aLa and SA, ammonium sulphate precipitation of BLG at pH 6	Dialysis Recrystallization
Armstrong, McKenzie & Sawyer (1961)	By ammonium sulphate precipitation at 20°C		Acid precipitation at pH 3.5 to remove aLa and SA, ammonium sulphate precipitation of BLG at pH 6	Dialysis Recrystallization
Monaco et al. (1987)	Centrifugation	Calcium chloride precipitation at pH 6.6	Dialysis, DEAE-cellulose chromatography using sodium chloride to elute BLG	Gel filtration Crystallization

Key
aLa = α-lactalbumin
SA = serum albumin

TABLE IV

QUANTITIES OF BLG IN THE MILK OF VARIOUS SPECIES.

Species	Quantity of BLG(mg/ml)	Reference
Bovine	1.8 to 5.0	[a]
Goat	1.4	[b]
Sheep	2.8	[c]
Red deer	2.8 to 3.0	[d]
Pig	0.6*	[e]
Dolphin	16.2	[f]
Manatee	14.1	[f]
Beagle	10.1	[f]
Human	0.000005-0.0008**	[g]
Mouse	0	[c]

Comments

* This value is an underestimate due to the poor fractionation of the porcine whey proteins at pH 3, and the numerous precipitation steps used subsequently.

** It is unlikely that this result refers to human BLG. Radioimmunoassays using antibovine BLG antisera have detected both fragments of lactoferrin and bovine BLG in human milk [Manti et al., 1989].

a Palmer, 1934; Butler, 1974
b Kalan & Basch, 1969
c Simons et al., 1987
d McDougall & Stewart, 1976
e Kalan et al., 1971
f Pervaiz & Brew, 1986
g Axelsson et al., 1986

TABLE V

AMINO ACID COMPOSITIONS OF PIG BLG

Amino Acid	Kalan[a] A	Kalan[a] B	Bell[b] A	Bell[b] C	Conti[c] I	Conti[d] 1	Conti[d] 2	Conti[d] 3	This work
Asx	17	17	17	18	17	16	18	18	17
Thr	10	10	10	9	10	10	9	10	10
Ser	10	10	9	9	8	7	9	7	7
Glx	23	23	24	22	24	22	21	24	24
Pro	8	8	8	8	7	7	6	7	8
Gly	3	3	4	4	3	7	5	4	3
Ala	13	14	14	13	13	13	12	13	13
Cys	4	4	4	4	5	4	6	5	nd
Val	13	12	13	13	13	15	10	12	12
Met	4	4	4	4	4	nd	2	3	3
Ile	6	6	6	6	6	6	5	6	6
Leu	24	24	25	24	24	22	27	25	25
Tyr	2	2	2	2	2	nd	5	3	2
Phe	3	3	3	3	3	3	6	3	3
Lys	11	11	11	11	11	14	11	11	12
His	3	3	3	4	3	4	4	4	4
Arg	5	5	5	5	5	5	4	5	6
Trp	1	1	1	1	1	nd	nd	nd	nd

Comments
nd = amino acid residue not dete
* = this amino acid composition was obtained from pig BLG, kindly supplied by Dr Carl Holt.

a Kalan et al., 1971
b Bell et al., 1981
c Conti et al., 1986a
d Conti et al., 1986b

TABLE VII

COMPARISON OF THE PRIMARY SEQUENCES OF NON-RUMINANT BLGS AND BOVINE BLG-A.

```
                 *       * *                    **          * *   **   *        *
Bovine A    LIVTQTMKGL DIQKVAGTWY SLAMAASDIS LLDAQSAPLR -VYVEELKPT PEGDLEI-LL QKWENDECAQ
Kangaroo    VENIRSKNDL GVEKFVGSWY LREAAKTMEF -----SIPLF DMDIKEVNLT PEGNLELVLL EKTDR--CVE
Dolphin I   VSVIRTMEDL DIQRVAGTWH SVAMAASDIS LLDTEEAPLR -VNVEELRPT PQGDLEL-FL QK.......
Donkey      TNIPQTMQDL DLQEVAGKWH SVAMAASDIS LLDSEEAPLR -VYIEKLRPT PEDNLEI-IL REGENKGCAE
Horse I     TNIPQTMQDL DLQEVAGKWH SVAMAASDIS LLDSEEAPLR -VYIEKLRPT PEDNLEI-IL REGENKGCAE
Horse II    TDIPQTMQDL DLQEVAGRWH SVAMVASDIS LLDSEEAPLR -VYVEELRPT PEGNLEI-IL REGANHACVE
Pig I       VEVTPIMTEL DTQKVAGTWH TVAMAVSDVS LLDAKESPLK -AYVEGLKPT PEGDLEI-LL CKRENDKCAQ

                ** *  * *                      *                              *
Bovine A    KKIIAEKTKI PAVFKIDALN ENKVL--VLD TDYKKYLLFC MENSAEPEQS ----LVCQCL VRTPEVDDEA
Kangaroo    KKLLLKKTKK PTEFEIYISS ESSYTFCVME TDYDSYFLFC LYNISDREK- ----MACAHY VRRIEENKGM
Dolphin I   .....EKTEI PAVF..NFLN .......... SDYTNYLLFC ME......VS ----LTCAYL ARTLQVDDGV
Donkey      KKIFAEKTES PAEFKINYLD EDTVF--ALD SDYKNYLFLC MKNAATPGQS ----LVCNYL ARTQMVDEEI
Horse I     KKIFAEKTQS PAQFKINALD EDTVF--YLD TDYKNYLFLC MKNAATPGQS ----LVC-YL ARTQMVDEEI
Horse II    RNIVAQKTES PAEFKTEDSA VFTVN--YQP GERKNYLFLC MKDVGPCLPS AEHGMVCQYL ARTQKVDEEV
Pig I       EVLLAKKTDI PAVFKINALD ENQLF--LLD TDYDSHLFLC MENASQEH-S ----LVQQCL ARTLEVDDQI

            * *                          *
Bovine A    LEKFDKALKA LPMHIRLSFN PTQLEEQCHI
Kangaroo    NE-FKKILRT LAMPYTV1-- EVRTRDMCHV
Dolphin I   MEKFNKAIKP ALPMHIR-FS PTQLEE
Donkey      MEKFRRALQP LPGRVQILPD LTRMAERCRI
Horse I     MEKFRRALQP LPGRVQIVPD LTRMA---RI
Horse II    MEKFSRALQP LPGRVQIVQD PSGGQERCGF
Pig I       REKFEDALKT LSVPMRIL-- PAQLEEQCRV
```

Key

.. residues not yet determined
- gap inserted to help alignment
* indicates a residue conserved in all the known sequences above

TABLE VI

COMPARISON OF THE AMINO ACID SEQUENCES OF RUMINANT BLGS.

	Variable amino acid positions																	
BLG	**1**	**11**	**20**	**28**	**45**	**50**	**53**	**59**	**64**	**78**	**84**	**87**	**118**	**129**	**130**	**148**	**150**	**158**	**162**
Bovine A	V	D	Y	D	E	P	D	Q	D	I	I	L	V	D	D	R	S	E	I
Bovine B	V	D	Y	D	E	P	D	Q	G	I	I	L	A	D	D	R	S	E	I
Bovine C	V	D	Y	D	E	P	D	H	G	I	I	L	A	D	D	R	S	E	I
Bovine D	V	D	Y	D	Q	P	D	Q	G	I	I	L	A	D	D	R	S	E	I
Bovine E	V	D	Y	D	E	P	D	Q	G	I	I	L	A	D	D	R	S	G	I
Bovine F	V	D	Y	D	E	S	D	Q	G	I	I	L	A	D/Y*		R	S	G	I
Bovine G	V	D	Y	D	E	P	D	Q	G	M	I	L	A	D	D	R	S	G	I
Bovine Dr	V	D	Y	N*	E	P	D	Q	D	I	I	L	V	D	D	R	S	E	I
Yak	V	N/D*	Y	D	E	P	D	Q	G	I	L	I	A	D	D	R	S	G	I
W'buffalo	I	D	Y	D	E	P	D	Q	D	I	I	L	A	D	D	R	S	E	V
Goat	I	D	Y	D	E	P	N	Q	D	I	I	L	A	D	K	R	A	G	V
Sheep A	I	D	Y	D	E	P	N	Q	D	I	I	L	A	D	N	R	A	G	V
Sheep B	I	D	H	D	E	P	N	Q	D	I	I	L	A	D	N	R	A	G	V
Sheep C	I	D	Y	D	E	P	N	Q	G	I	I	L	A	D	N	Q	A	G	V

Key

D/Y* : It is unclear which residue is at which position since the changes in E, F and G are inferred (Bell et al., 1981)

N/D* : This position was originally believed to be Asn (Grosclaude et al., 1976), but has also been reported as Asp (Bell et al., 1981)

W'buffalo : Waterbuffalo

TABLE VIII - Selected physicochemical parameters for ß-lactoglobulin

Parameter	Value	Reference
Isoelectric point	5.18	Cannan *et al* (1942)
Molecular mass (Daltons)	36,000	Townend & Timasheff (1957)
Sedimentation coefficient ($S°_{20w} \times 10^{13}\ s^{-1}$)	2.83	Cecil & Ogston (1949)
Diffusion coefficient ($cm^2\ sec^{-1} \times 10^{-7}$)	7.70 ± 0.09	Cecil & Ogston (1949)
Crystal density for salt-free Lattice Q ($g.cm^{-3}$)	1.147 ± 0.003	Green *et al.* (1956)
Intrinsic viscosity [η] (ml/g)	2.9	Townend *et al* (1960b)
Partial specific volume (cm^3/g)	0.751	Swaisgood (1982)
Hydration (%)	35 - 40	Pantaloni (1965)
Axial ratio	2:1	Green & Aschaffenburg (1959)
Stokes radius (nm)	2.68	Swaisgood (1982)
Dipole moment (Debye)	730	Ferry & Oncley (1941)
Refractive Index (at 589nm)	1.594	Swaisgood (1982)
R_G (nm)	2.17	Witz *et al* (1964)
Extinction coefficient ($l\ g^{-1}\ cm^{-1}$)	0.96	Townend *et al* (1960b)
Charge at pH 6.6	-10.0	Basch & Timasheff (1967)

TABLE IX

PARAMETERS FOR LIGAND BINDING TO BOVINE BLG.

Ligand	Number bound per dimer	Association constant (M^{-1})	Reference
Retinol	2	5×10^7	[a]
Stearate	2	1.7×10^5	[b]
Palmitate	2	6.8×10^5	[b]
Laurate	2	0.5×10^5	[b]
Oleate	2	0.4×10^5	[b]
Heptane	2	0.48×10^6	[c]
Butane	2	1.7×10^3	[c]
Pentane	2	(5.8×10^2) 7.1×10^3	[d]
Iodobutane	2	(6.2×10^3) 2.8×10^3	[d]
SDS	2	3.1×10^5	[e]
2,6-MANS	2	0.2×10^4	[f]
Methyl orange	3	6.3×10^4	[e]
n-Octylbenzene-p-sulphonate	2	1.9×10^4	[g]
p-Nitrophenol	2	3.0×10^4	[g]
p-Nitrophenylacetate	2	1.6×10^4	[g]
p-Nitrophenyl-B-glucuronide	2	2.0×10^3	[g]
p-Nitrophenyl sulphate	2	3.1×10^3	[g]
p-Nitrophenyl pyridoxal phosphate	2	0.15×10^3	[e]
2-Heptanone	2	0.50×10^3	[e]
2-Octanone	2	2.44×10^3	[e]
2-Nonanone	2	4.5×10^2 (5.9×10^1)	[h]
Toluene	2	4.2×10^2 (3.1×10^1)	[h]
Trifluorotoluene	2	1.6×10^3	[h]
Hexafluorobenzene			

Comments
Figures in brackets represent the association constant for the binding of a second ligand at the same site.

a Fugate & Song (1980)
b Spector & Fletcher (1970)
c Mohammed-K et al. (1969)
d Wishnia & Pinder (1966)
e O'Neill & Kinsella (1987)
f Lovrien & Anderson (1969)
g Farrell et al. (1987)
h Robillard & Wishnia (1972)

TABLE X - SOME CHARACTERISTICS OF THE BLG-SUPERFAMILY PROTEINS

Protein	Number of Residues	Location	Ligand	Function	Ref.
Bovine BLG	162	Milk	Retinol, FFA?	Transport/transfer in gut of young?	a
PP14 (α-PEG)	162	Amniotic fluid	-	-	b,c
Human RBP	182	Serum	Retinol	Retinol transport in complex with transthyretin	d
Purpurin	175	Neural retina	Retinol	Transport retinol across interphotoreceptor matrix	e
α_1-Microglobulin	167	Serum, urine,	Brown-yellow	Binds IgA or albumin	f
Protein HC	181	cerebrospinal fluid	chromophore		g
Frog BG protein	159	Nasal mucosa	Pyrazines, other odorants	Presentation of odorants to receptors on neurons	h
Cow nasal protein	159	"			i
Odorant BP	172	"			j
Aphrodisin	151	Vaginal discharge	-	Aphrodisiac	k
Apolipoprotein-D	169	Serum	Cholesteryl esters?	Part of LCAT complex? Transports cholesterol or bilin?	l, m
α_1-acid glycoprotein	181	Serum, urine	Steroids, unknown lipid	Acute phase protein	n
Rat ESP	166	Epididymal luminal fluid	-	Binds to sperm membrane, and aids maturation	o
α_{2u}-globulin	162	Urine, serum	Pheromones	-	p
MUP	162	"	"	-	q
Complement C8γ	182	Serum	.....	Inessential bit of complement complex that lyses cells	r
Crustacyanin A	168	Carapace	Astaxanthin	Colouration	s
Crustacyanin C	175				
Insecticyanin	189	Hemolymph	Biliverdin IX	Camouflage	t
Bilin BP	173				u

Key to Table X

α-PEG	Pregnancy-associated endometrial α-globulin
BG	Bowman's gland protein
BP	Binding protein
ESP	Epididymal protein
MUP	Mouse major urinary protein

a	Godovac-Zimmermann & Braunitzer (1987)
b	Julkunen et al. (1988)
c	Rask *et al.* (1979)
d	Berman *et al.* (1987)
e	Åkerström & Lögdberg (1990)
f	Pervaiz & Brew (1985)
h	Lee *et al.* (1987)
i	Tirindelli *et al.* (1989)
j	Pevsner *et al.* (1988)
k	Henzel *et al.* (1988)
l	Drayna *et al.* (1986)
m	Peitsch & Boguski (1990)
n	Cooper *et al.* (1987)
o	Brooks *et al.* (1986)
p	Unterman *et al.* (1981)
q	Clark *et al.* (1984)
r	Haefliger *et al.* (1987)
s	Keen *et al.* (1990)
t	Riley *et al.* (1984)
u	Sutor *et al.* (1988)

FIGURE LEGENDS

Figure 1

Specific rotation as a function of pH showing the conformational changes for bovine ß-lactoglobulin variants A, B and C (data from McKenzie & Sawyer, 1967) and for pig (■ - P, S.G.Hambling, unpublished results). The curve S corresponds to the denatured state (Schellman, 1958).

Figure 2

a A cartoon representation of the backbone together with the strand identification of bovine ß-lactoglobulin, based upon a Figure from Papiz *et al.* (1986).

b A stereodiagram of the backbone trace of the lattice Y crystal form of bovine ß-lactoglobulin showing the positions of the disulphide bridge 106-119 and and the free thiol at 121, the Trp-19 and the Tyr side chains. The view has the entrance to the calyx on the left and the helix-sheet putative binding site in the centtre around Cys-121.

c A stereodiagram showing the dimer looking approximately along the twofold axis of symmetry. The subunits are arranged one on top of the other on the page.

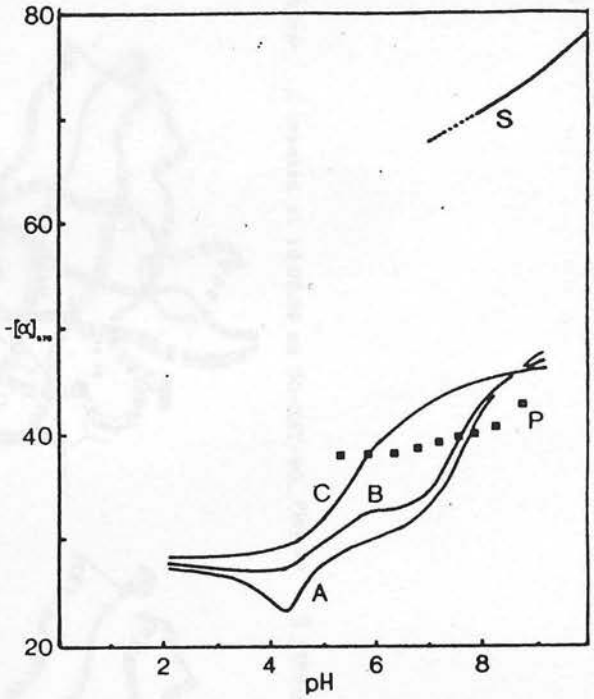


Figure 1

Picture 2 created at 18:06:50 on 30-MAY-90 Picture 3 created at 18:07:34 on 30-MAY-90

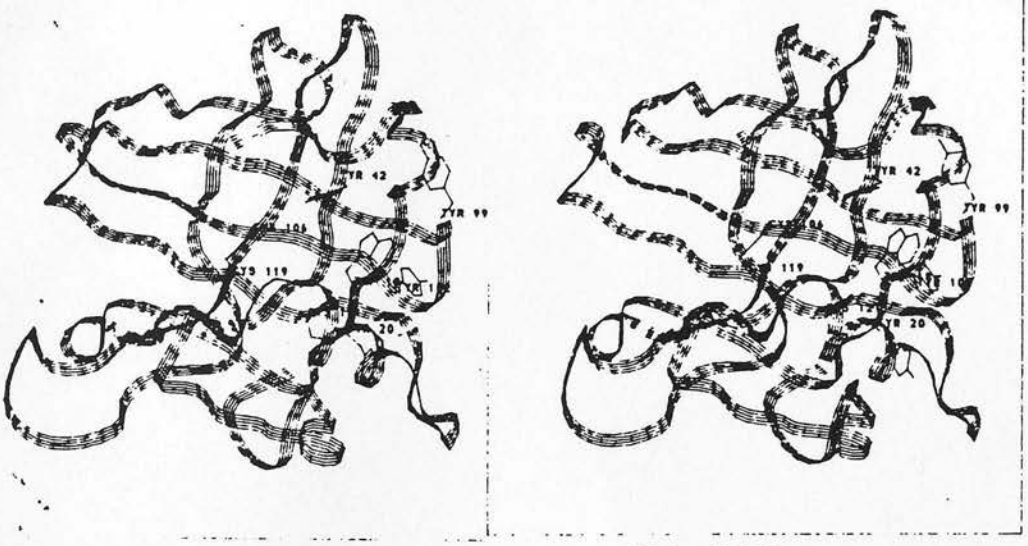


fig 2a